GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 5, 2005, 14:55:35 ; Search time 47.8306 Seconds
(without alignments)
1245.251 Million cell updates/sec

Title: US-10-090-185-8
Perfect score: 800
Sequence: 1 MAQWNQLQQLDTRYLKQLHQ.....AVVTEKQQMLEQHLQDVRKR 154

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Search Notes

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	800	100.0	154	4	AAY72840	Aay72840 Mouse Sta
2	800	100.0	770	2	AAR72082	Aar72082 Mouse Sta
3	800	100.0	770	2	AAW03176	Aaw03176 Mouse STA
4	800	100.0	770	3	AAB12377	Aab12377 N-termina
5	800	100.0	770	5	AAE14652	Aae14652 Murine ST
6	800	100.0	770	6	ABU10476	Abu10476 Mouse STA
7	792	99.0	720	5	AAE22055	Aae22055 Human Sta
8	792	99.0	769	5	ABB57164	Abb57164 Mouse isc
9	792	99.0	769	5	AAE22054	Aae22054 Human Sta
10	792	99.0	769	5	AAE22056	Aae22056 Human pro
11	792	99.0	770	2	AAR82995	Aar82995 Mouse liv
12	792	99.0	770	2	AAR82993	Aar82993 Human pla
13	792	99.0	770	2	AAY03768	Aay03768 Human STA
14	792	99.0	770	4	AAB19964	Aab19964 Human sig
15	792	99.0	770	5	ABG69497	Abg69497 Human bai
16	792	99.0	770	5	AAE15174	Aae15174 Human Sta
17	792	99.0	770	7	ADD44738	Add44738 Rat Prote
18	792	99.0	770	7	ADD44740	Add44740 Human Pro
19	792	99.0	770	8	ADN04365	Adn04365 Antipsori
20	792	99.0	770	8	ADP54789	Adp54789 Human PRO
21	792	99.0	793	3	AAB58442	Aab58442 Lung canc
22	380	47.5	154	4	AAY72843	Aay72843 Mouse Sta
23	380	47.5	680	6	ABR59713	Abr59713 Human sig
24	380	47.5	701	2	AAR41335	Aar41335 84 kD ISG
25	380	47.5	701	6	ABU04741	Abu04741 Human exp
26	380	47.5	712	2	AAR72079	Aar72079 Human Sta
27	380	47.5	712	2	AAW03170	Aaw03170 Human STA
28	380	47.5	712	2	AAW62995	Aaw62995 Human Sta
29	380	47.5	712	6	ABU04747	Abu04747 Human exp
30	380	47.5	712	6	ABU04735	Abu04735 Human exp
31	380	47.5	712	6	ABU04745	Abu04745 Human exp
32	380	47.5	712	6	ABU04743	Abu04743 Human exp
33	380	47.5	712	8	ADH57035	Adh57035 Truncated
34	380	47.5	712	8	ADJ32352	Adj32352 Human STA
35	380	47.5	712	8	ADN04464	Adn04464 Antipsori
36	380	47.5	729	5	AAU84356	Aau84356 Protein S
37	380	47.5	739	2	AAR41334	Aar41334 91 kD ISG
38	380	47.5	739	6	ABU04740	Abu04740 Human exp
39	380	47.5	750	2	AAR72078	Aar72078 Human Sta
40	380	47.5	750	2	AAW03168	Aaw03168 Human STA
41	380	47.5	750	2	AAW62994	Aaw62994 Human Sta
42	380	47.5	750	4	AAB19962	Aab19962 Human sig
43	380	47.5	750	5	AAE15172	Aae15172 Human Sta
44	380	47.5	750	5	AAG78525	Aag78525 Human STA
45	380	47.5	750	6	ABP98871	Abp98871 Human STA

ALIGNMENTS

RESULT 1
AAY72840
ID AAY72840 standard; protein; 154 AA.
XX
AC AAY72840;
XX
DT 31-MAY-2001 (first entry)
XX
DE Mouse Stat3 protein fragment #1 (1-154 amino acids).
XX
KW Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW cellular transformation; dysproliferative disease; cancer; psoriasis;
KW therapy.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT Domain 1. .130
FT /note= "N-terminal domain"
FT Region 130. .154
FT /note= "Stat3-c-Jun interaction region 1"
XX
PN WO200116605-A2.
XX
PD 08-MAR-2001.
XX
PF 30-AUG-2000; 2000WO-US023822.
XX
PR 31-AUG-1999; 99US-00387418.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI Zhang X, Horvath C, Wrzeszcynska MH, Darnell JE;
XX
DR WPI; 2001-226705/23.
XX
PT Identifying an agent for use in modulating the interaction between
PT transcription factor c-Jun and a Stat3 protein.
XX
PS Claim 65; Page 66-67; 86pp; English.
XX
CC The present sequence is mouse Stat3 protein fragment containing 1-154
CC amino acids of Stat3 protein. This Stat3 fragment showed very weak
CC binding to c-Jun protein in the cell extract. The invention relates to
CC methods for identifying interacting regions of transcription factors and
CC methods for identifying agents which modulates the interaction between a
CC transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC Stat-3, useful for modulating gene transcription e.g., cellular

```
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 154 AA;

  Query Match           100.0%;  Score 800;  DB 4;  Length 154;
  Best Local Similarity  100.0%;  Pred. No. 1.1e-77;
  Matches  154;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 2
AAR72082
ID   AAR72082 standard; protein; 770 AA.
XX
AC   AAR72082;
XX
DT   25-MAR-2003  (revised)
DT   27-SEP-1995  (first entry)
XX
DE   Mouse Stat3 (19sf6).
XX
KW   Signal transducer and activator of transcription; STAT; 19sf6; Stat3;
KW   receptor recognition factor; transcription factor; cellular debilitation;
KW   derangement; dysfunction; interferon-gamma.
XX
OS   Mus sp.
XX
PN   WO9508629-A1.
XX
PD   30-MAR-1995.
XX
PF   26-SEP-1994;   94WO-US010849.
XX
PR   24-SEP-1993;   93US-00126588.
PR   24-SEP-1993;   93US-00126595.
PR   11-MAR-1994;   94US-00212184.
PR   11-MAR-1994;   94US-00212185.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Darnell JE,  Schindler CW,  Shuai K,  Wen Z,  Zhong Z;
XX
DR   WPI; 1995-139598/18.
DR   N-PSDB; AAQ89340.
XX
PT   Receptor recognition factor implicated in transcriptional stimulation of
PT   genes - useful in drug screening assays and/or for treating cellular
PT   debilitations, derangements and/or dysfunctions, etc.
XX
PS   Claim 1; Page 107-110; 160pp; English.
XX
CC   A fragment encoding the human Stat91 protein was used to screen a murine
CC   thymus and spleen cDNA for homologous proteins. A highly homologous gene
CC   (given in AAQ89338) was isolated that encoded a 91 kDa protein (AAR72080)
CC   (Stat1) that was responsive to interferon- gamma. Using a fragment of the
CC   mouse gene as probe, 2 additional members of the 113-91 family of
CC   receptor recognition factor proteins were isolated. The 2 genes (AAQ89339
CC   -40) were cloned in plasmids 13sf1 and 19sf6 and encoded proteins termed
CC   Stat4 (AAR72081) and Stat3 (AAR72082), respectively. (Updated on 25-MAR-
CC   2003 to correct PN field.)
XX
SQ   Sequence 770 AA;

  Query Match           100.0%;  Score 800;  DB 2;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 9e-77;
  Matches  154;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 3
AAW03176
ID   AAW03176 standard; protein; 770 AA.
XX
AC   AAW03176;
XX
DT   24-OCT-1996  (first entry)
XX
DE   Mouse STAT4.
XX
KW   STAT; STAT4; signal transducer and activator of transcription;
KW   DNA binding protein; ligand; receptor; oncogenesis; inflammation;
KW   autoimmune disease; antagonist; therapy.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Domain          398. .508
FT                   /label= DNA_binding_domain
FT                   /note= "Claim 3, page 110"
XX
PN   WO9620954-A2.
XX
PD   11-JUL-1996.
XX
PF   28-DEC-1995;   95WO-US017025.
XX
PR   06-JAN-1995;   95US-00369796.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Darnell JE,  Wen Z,  Horvath CM,  Zhong Z;
XX
DR   WPI; 1996-333941/33.
DR   N-PSDB; AAT31280.
XX
PT   New STAT protein DNA-binding domain peptide(s) - useful for diagnosing,
PT   preventing or treating cellular dysfunction, e.g. oncogenesis,
PT   inflammation, parasitic disease or autoimmunity.
XX
PS   Disclosure; Page 87-90; 138pp; English.
XX
CC   Mouse signal transducer and activator of transcription (STAT) protein
CC   STAT4 (AAW03176) serves a dual purpose, i.e. signal transduction from
CC   ligand-activated receptor kinase complexes followed by nuclear
CC   translocation and DNA binding to activate transcription. Recombinant
CC   STAT4 can be obtd. using cDNA clone 19sf6 (AAT31278) obtd. from
CC   splenic/thymic cells. STAT4 includes a DNA-binding domain (see also
CC   AAW03167) capable of both receptor recognition and message delivery via
CC   DNA binding in a receptor-ligand specific manner. STAT proteins and their
CC   DNA binding domains (see also AAW03165-75) are useful for screening
CC   antagonists used to inhibit STAT-mediated signal transduction and
```

```
CC   activation of transcription
XX
SQ   Sequence 770 AA;

  Query Match             100.0%;  Score 800;  DB 2;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 9e-77;
  Matches  154;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 4
AAB12377
ID   AAB12377 standard; peptide; 770 AA.
XX
AC   AAB12377;
XX
DT   08-NOV-2000  (first entry)
XX
DE   N-terminal domain of murine STAT-3 protein.
XX
KW   STAT; signal transducer and activator of transcription; crystal;
KW   drug design; murine.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Region          4. .9
FT                   /label= Alpha helix 1
FT   Region          12. .21
FT                   /label= Alpha helix 2
FT   Region          19. .21
FT                   /label= 3(10) helix of alpha helix 2
FT   Region          28. .33
FT                   /label= Alpha helix 3
FT   Region          35. .40
FT                   /label= Alpha helix 4
FT   Region          43. .47
FT                   /label= Alpha helix 5
FT   Region          50. .73
FT                   /label= Alpha helix 6
FT   Region          77. .96
FT                   /label= Alpha helix 7
FT   Region          99. .119
FT                   /label= Alpha helix 8
XX
PN   US6087478-A.
XX
PD   11-JUL-2000.
XX
PF   23-JAN-1998;   98US-00012710.
XX
PR   23-JAN-1998;   98US-00012710.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Vinkemeier U,  Moarefi I,  Darnell JE,  Kuriyan J;
XX
DR   WPI; 2000-505108/45.
XX
PT   New crystals of an N-terminal fragment of a signal transducer and
PT   activator of transcription that effectively diffracts x-rays, useful for
PT   drug screening and development.
XX
PS   Disclosure; Fig 1; 42pp; English.
XX
CC   The present invention relates to a crystal of an N-terminal fragment of a
CC   signal transducer and activator of transcription (STAT) protein. The
CC   crystal effectively diffracts X-rays, allowing the determination of the
CC   atomic coordinates of the N-terminal domain to a resolution of greater
CC   than 5.0 Angstroms. The present sequence is the N-terminal domain of the
CC   murine STAT 3 protein. The N-terminal domain enables STAT dimers to
CC   interact and bind DNA cooperatively, a mechanism important for gene
CC   activation. The crystals are useful in drug screening and development by
CC   selecting a potential drug by performing rational drug design with the 3-
CC   dimensional structure determined for the crystal
XX
SQ   Sequence 770 AA;

  Query Match             100.0%;  Score 800;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 9e-77;
  Matches  154;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 5
AAE14652
ID   AAE14652 standard; protein; 770 AA.
XX
AC   AAE14652;
XX
DT   16-JUL-2002  (first entry)
XX
DE   Murine STAT3 protein.
XX
KW   Signal transducer and activator of transcription; STAT3;
KW   drug development; drug discovery; crystal; inflammation; allergy; asthma;
KW   leukaemia; anaemia; neutropaenia; thrombocytopaenia; cancer; obesity;
KW   viral disease; growth retardation; murine.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Domain          1. .130
FT                   /note= "Conserved N-terminal domain of the STAT family"
XX
PN   US6312887-B1.
XX
PD   06-NOV-2001.
XX
PF   24-APR-2000; 2000US-00556273.
XX
PR   23-JAN-1998;   98US-00012710.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Vinkemeier U,  Moarefi I,  Darnell JE,  Kuriyan J;
XX
DR   WPI; 2002-033337/04.
XX
PT   Identifying compounds that bind to signal transducer and activator of
PT   transcription proteins, useful for the production of new drugs.
XX
PS   Example; Col 47-50; 44pp; English.
XX
```

```
CC   The invention relates to methods for detecting compounds that bind to
CC   signal transducer and activator of transcription (STAT) proteins for the
CC   discovery and development of new drug compounds based on the structural
CC   properties of the protein crystal. The methods include: identifying a
CC   compound that binds to the N-terminal domain of a STAT protein,
CC   identifying a compound that enhances or diminishes the binding of the
CC   dimeric STAT proteins to each other and/or their nucleic acid binding
CC   site; or identifying a compound that enhances or diminishes the ability
CC   of STAT protein dimers to induce the expression of a gene operably under
CC   the control of a promoter containing at least two adjacent weak binding
CC   sites for STAT protein dimers. The methods are used for identifying new
CC   drugs. An antagonist of STAT N-terminal dimeric interactions that
CC   inhibits the binding of the STAT dimers to adjacent weak binding sites on
CC   a promoter of a gene, could be useful as drugs in the treatment of
CC   diseases, e.g. inflammation, allergy, asthma and leukaemias. On the other
CC   hand, an agonist of N-terminal dimeric interactions between STAT dimers,
CC   can be used as drugs in the treatment of diseases e.g. anaemia,
CC   neutropaenia, thrombocytopaenia, cancer, obesity, viral diseases and
CC   growth retardation. The present sequence is murine STAT3 protein
XX
SQ   Sequence 770 AA;

  Query Match             100.0%;  Score 800;  DB 5;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 9e-77;
  Matches  154;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 6
ABU10476
ID   ABU10476 standard; protein; 770 AA.
XX
AC   ABU10476;
XX
DT   06-AUG-2003  (first entry)
XX
DE   Mouse STAT3 protein..
XX
KW   Mouse; signal transducer and activator of transcription; drug design;
KW   drug screening; STAT-STAT dimer interaction; STAT3.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Region          4. .9
FT                   /label= alpha_helix_1
FT   Region          12. .21
FT                   /label= alpha_helix_2
FT                   /note= "Residues 19-21 form a 3 helix"
FT   Region          28. .33
FT                   /label= alpha_helix_3
FT   Region          35. .40
FT                   /label= alpha_helix_4
FT   Region          43. .47
FT                   /label= alpha_helix_5
FT   Region          50. .73
FT                   /label= alpha_helix_6
FT                   /note= "Residues 57, 61, 64, 68 and 71 contribute to
FT                   packing of the coiled-coil"
FT   Region          77. .96
FT                   /label= alpha_helix_7
FT                   /note= "Residues 79, 83, 86, 90 and 94 contribute to
FT                   packing of the coiled-coil"
FT   Region          99. .119
FT                   /label= alpha_helix_8
XX
PN   US2003003563-A1.
XX
PD   02-JAN-2003.
XX
PF   19-OCT-2001; 2001US-00045792.
XX
PR   23-JAN-1998;   98US-00012710.
PR   24-APR-2000; 2000US-00556273.
XX
PA   (VINK/) VINKEMEIER U.
PA   (MOAR/) MOAREFI I.
PA   (DARN/) DARNELL J E.
PA   (KURI/) KURIYAN J.
XX
PI   Vinkemeier U,  Moarefi I,  Darnell JE,  Kuriyan J;
XX
DR   WPI; 2003-447354/42.
XX
PT   New crystal having an N-terminal domain of a STAT protein performing X-
PT   ray crystallographic studies, useful for screening drugs that enhance or
PT   inhibit STAT-STAT dimer interactions.
XX
PS   Disclosure; Page 25-26; 46pp; English.
XX
CC   The invention relates to a crystal of an N-terminal domain of signal
CC   transducer and activator of transcription (STAT) protein, where the
CC   crystal effectively diffracts X-rays for the determination of the atomic
CC   coordinates of the N-terminal domain of the STAT protein to a resolution
CC   of greater than 5.0 Angstrom. The methods and compositions are useful for
CC   the design and screening of drugs that enhance or inhibit STAT-STAT dimer
CC   interactions. The present sequence represents the amino acid sequence of
CC   mouse STAT3 protein
XX
SQ   Sequence 770 AA;

  Query Match             100.0%;  Score 800;  DB 6;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 9e-77;
  Matches  154;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 7
AAE22055
ID   AAE22055 standard; protein; 720 AA.
XX
AC   AAE22055;
XX
DT   25-JUL-2002  (first entry)
XX
DE   Human Stat3beta protein.
XX
KW   Human; signal transducer and activator of transcription 3; ischaemia;
KW   immune response; Stat3; coronary atherosclerosis; vascular occlusion;
KW   hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;
KW   inflammation; chronic obstructive pulmonary disease; cardiac arrest;
KW   insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;
KW   shock; chronic active hepatitis; adult respiratory distress syndrome;
KW   nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;
```

KW Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;
KW polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;
KW proliferative disease; Grave's disease; ulcerative colitis; sarcoma;
KW carcinoma; degenerative disorder; gene therapy; growth deficiency;
KW cirrhosis; hypoproliferative disorder; lesion; Statbeta.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 713. .714
FT /note= "Encoded by ACA CCA TTC"
XX
PN WO200220032-A1.
XX
PD 14-MAR-2002.
XX
PF 10-SEP-2001; 2001WO-US028254.
XX
PR 08-SEP-2000; 2000US-0231212P.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
PA (UYSF-) UNIV SOUTH FLORIDA.
XX
PI Yu H, Pardoll D, Jove R, Dalton W;
XX
DR WPI; 2002-362218/39.
DR N-PSDB; AAD35066.
XX
PT Modulating angiogenesis and an immune response in an individual, for
PT treating a hypoxic or ischemic condition, comprises administering a
PT compound that modulates the activity of a signal transducer and activator
PT of transcription 3.
XX
PS Disclosure; Page 87-89; 94pp; English.
XX
CC The invention relates to a method of modulating angiogenesis and immune
CC response. Method involves administering to an individual a compound that
CC modulate the activity of signal transducer and activator of transcription
CC 3 (Stat3). Modulating angiogenesis is useful for treating or preventing
CC hypoxic or ischaemic condition or disorder which is the result of stroke,
CC ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,
CC tissue ischaemia in the lower extremities, infarction, trauma, vascular
CC occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,
CC chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,
CC epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest,
CC nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy
CC with neovascularisation. Suppressing an immune response is useful for
CC ameliorating a symptom of an autoimmune disease such as systemic lupus
CC erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,
CC Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,
CC mixed connective tissue disease, primary biliary cirrhosis, pernicious
CC anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,
CC gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,
CC idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's
CC disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune
CC infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and
CC dense deposit disease. The method is useful in preventing or treating
CC specific proliferative and oncogenic disease which includes sarcomas and
CC carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,
CC fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,
CC hypoproliferative disorders, physical trauma, lesions and wounds. The
CC method is also used in gene therapy. The present sequence is human
CC Stat3beta protein
XX
SQ Sequence 720 AA;

Query Match 99.0%; Score 792; DB 5; Length 720;
Best Local Similarity 98.7%; Pred. No. 6e-76;
Matches 152; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
||||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db 1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy 61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy 121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
||||||||||||||||||||||||||||||||
Db 121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154

RESULT 8
ABB57164
ID ABB57164 standard; protein; 769 AA.
XX
AC ABB57164;
XX
DT 07-MAR-2002 (first entry)
XX
DE Mouse ischaemic condition related protein sequence SEQ ID NO:398.
XX
KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease.
XX
OS Mus musculus.
XX
PN WO200188188-A2.
XX
PD 22-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-JP004192.
XX
PR 18-MAY-2000; 2000JP-00145977.
XX
PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
DR WPI; 2002-034733/04.
DR N-PSDB; ABI99454.
XX
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or by
PT determining the expression profile of a gene group comprising these
PT genes.
XX
PS Claim 2; Page 1084-1087; 2690pp; English.
XX
CC The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (ABI99202 to ABI99912, encoding the
CC protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The expression
CC levels or expression profiles produced by these genes are used as an
CC indicator when screening for ischaemic condition-improving drugs or
CC therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR
CC primers for a mouse ischaemic condition related sequence, which are used
CC in the exemplification of the present invention
XX
SQ Sequence 769 AA;

Query Match 99.0%; Score 792; DB 5; Length 769;
Best Local Similarity 98.7%; Pred. No. 6.5e-76;
Matches 152; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
||||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db 1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy 61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy         121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
                |||||||||||||||||||||||||||||||||||
Db         121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
```

RESULT 9
AAE22054
ID AAE22054 standard; protein; 769 AA.
XX
AC AAE22054;
XX
DT 25-JUL-2002 (first entry)
XX
DE Human Stat3 protein.
XX
KW Human; signal transducer and activator of transcription 3; ischaemia;
KW immune response; Stat3; coronary atherosclerosis; vascular occlusion;
KW hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;
KW inflammation; chronic obstructive pulmonary disease; cardiac arrest;
KW insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;
KW shock; chronic active hepatitis; adult respiratory distress syndrome;
KW nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;
KW Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;
KW polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;
KW proliferative disease; Grave's disease; ulcerative colitis; sarcoma;
KW carcinoma; degenerative disorder; gene therapy; growth deficiency;
KW cirrhosis; hypoproliferative disorder; lesion.
XX
OS Homo sapiens.
XX
PN WO200220032-A1.
XX
PD 14-MAR-2002.
XX
PF 10-SEP-2001; 2001WO-US028254.
XX
PR 08-SEP-2000; 2000US-0231212P.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
PA (UYSF-) UNIV SOUTH FLORIDA.
XX
PI Yu H, Pardoll D, Jove R, Dalton W;
XX
DR WPI; 2002-362218/39.
DR N-PSDB; AAD35065.
XX
PT Modulating angiogenesis and an immune response in an individual, for
PT treating a hypoxic or ischemic condition, comprises administering a
PT compound that modulates the activity of a signal transducer and activator
PT of transcription 3.
XX
PS Disclosure; Page 83-85; 94pp; English.
XX
CC The invention relates to a method of modulating angiogenesis and immune
CC response. Method involves administering to an individual a compound that
CC modulate the activity of signal transducer and activator of transcription
CC 3 (Stat3). Modulating angiogenesis is useful for treating or preventing
CC hypoxic or ischaemic condition or disorder which is the result of stroke,
CC ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,
CC tissue ischaemia in the lower extremities, infarction, trauma, vascular
CC occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,
CC chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,
CC epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest,
CC nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy
CC with neovascularisation. Suppressing an immune response is useful for
CC ameliorating a symptom of an autoimmune disease such as systemic lupus
CC erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,
CC Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,
CC mixed connective tissue disease, primary biliary cirrhosis, pernicious
CC anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,
CC gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,
CC idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's
CC disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune
CC infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and
CC dense deposit disease. The method is useful in preventing or treating
CC specific proliferative and oncogenic disease which includes sarcomas and
CC carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,
CC fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,
CC hypoproliferative disorders, physical trauma, lesions and wounds. The
CC method is also used in gene therapy. The present sequence is human Stat3
CC protein
XX
SQ Sequence 769 AA;

```
  Query Match             99.0%;  Score 792;  DB 5;  Length 769;
  Best Local Similarity   98.7%;  Pred. No. 6.5e-76;
  Matches  152;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              |||||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
```

RESULT 10
AAE22056
ID AAE22056 standard; protein; 769 AA.
XX
AC AAE22056;
XX
DT 25-JUL-2002 (first entry)
XX
DE Human protein related to angiogenesis regulation.
XX
KW Human; signal transducer and activator of transcription 3; ischaemia;
KW immune response; Stat3; coronary atherosclerosis; vascular occlusion;
KW hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;
KW inflammation; chronic obstructive pulmonary disease; cardiac arrest;
KW insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;
KW shock; chronic active hepatitis; adult respiratory distress syndrome;
KW nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;
KW Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;
KW polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;
KW proliferative disease; Grave's disease; ulcerative colitis; sarcoma;
KW carcinoma; degenerative disorder; gene therapy; growth deficiency;
KW cirrhosis; hypoproliferative disorder; lesion.
XX
OS Homo sapiens.
XX
PN WO200220032-A1.
XX
PD 14-MAR-2002.
XX
PF 10-SEP-2001; 2001WO-US028254.
XX
PR 08-SEP-2000; 2000US-0231212P.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
PA (UYSF-) UNIV SOUTH FLORIDA.
XX
PI Yu H, Pardoll D, Jove R, Dalton W;
XX
DR WPI; 2002-362218/39.
XX
PT Modulating angiogenesis and an immune response in an individual, for
PT treating a hypoxic or ischemic condition, comprises administering a

PT compound that modulates the activity of a signal transducer and activator
PT of transcription 3.
XX
PS Disclosure; Page 83-85; 94pp; English.
XX
CC The invention relates to a method of modulating angiogenesis and immune
CC response. Method involves administering to an individual a compound that
CC modulate the activity of signal transducer and activator of transcription
CC 3 (Stat3). Modulating angiogenesis is useful for treating or preventing
CC hypoxic or ischaemic condition or disorder which is the result of stroke,
CC ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,
CC tissue ischaemia in the lower extremities, infarction, trauma, vascular
CC occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,
CC chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,
CC epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest,
CC nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy
CC with neovascularisation. Suppressing an immune response is useful for
CC ameliorating a symptom of an autoimmune disease such as systemic lupus
CC erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,
CC Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,
CC mixed connective tissue disease, primary biliary cirrhosis, pernicious
CC anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,
CC gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,
CC idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's
CC disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune
CC infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and
CC dense deposit disease. The method is useful in preventing or treating
CC specific proliferative and oncogenic disease which includes sarcomas and
CC carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,
CC fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,
CC hypoproliferative disorders, physical trauma, lesions and wounds. The
CC method is also used in gene therapy. The present sequence is human
CC protein related to angiogenesis regulation
XX
SQ Sequence 769 AA;

Query Match 99.0%; Score 792; DB 5; Length 769;
Best Local Similarity 98.7%; Pred. No. 6.5e-76;
Matches 152; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              |||||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
```

RESULT 11
AAR82995
ID AAR82995 standard; protein; 770 AA.
XX
AC AAR82995;
XX
DT 25-MAR-1996 (first entry)
XX
DE Mouse liver acute phase response factor.
XX
KW Mouse; acute phase response factor; transcription factor; interleukin-6;
KW signal transmission; liver; antibody; antisense; ribozyme;
KW antiinflammatory; antitumor; hypotensive; therapy.
XX
OS Mus musculus.
XX
PN EP676469-A2.
XX
PD 11-OCT-1995.
XX
PF 29-MAR-1995; 95EP-00104670.
XX
PR 04-APR-1994; 94JP-00065825.
XX
PA (KISH/) KISHIMOTO T.
XX
PI Akira S, Kishimoto T;
XX
DR WPI; 1995-346089/45.
DR N-PSDB; AAT05619.
XX
PT New acute phase response factor - for developing inhibitory agents for
PT treating diseases induced by cytokine(s) such as IL-6, e.g. inflammatory
PT diseases.
XX
PS Claim 10; Page 20-22; 31pp; English.
XX
CC The sequence represents a mouse acute phase response factor (APRF), a
CC transcription factor related to signal transmission of interleukin-6 (IL-
CC 6). The protein is encoded by a cDNA, isolated from a mouse liver cDNA
CC library using a polymerase chain reaction product (amplified using
CC primers derived from an IL-6-treated mouse liver peptide) as a probe.
CC APRF-inhibitors, e.g. antibodies, antisense oligonucleotides or
CC ribozymes, may be used to treat diseases induced by IL-6, e.g.
CC inflammatory disease, leukemia, cancer, osteoclasia, pulmonary
CC hypertension, etc
XX
SQ Sequence 770 AA;

Query Match 99.0%; Score 792; DB 2; Length 770;
Best Local Similarity 98.7%; Pred. No. 6.5e-76;
Matches 152; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              |||||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
```

RESULT 12
AAR82993
ID AAR82993 standard; protein; 770 AA.
XX
AC AAR82993;
XX
DT 25-MAR-1996 (first entry)
XX
DE Human placenta acute phase response factor protein.
XX
KW human; acute phase response factor; transcription factor; interleukin-6;
KW signal transmission; placenta; antibody; antisense; ribozyme;
KW antiinflammatory; antitumor; hypotensive; therapy.
XX
OS Homo sapiens.
XX
PN EP676469-A2.
XX
PD 11-OCT-1995.
XX
PF 29-MAR-1995; 95EP-00104670.
XX
PR 04-APR-1994; 94JP-00065825.
XX
PA (KISH/) KISHIMOTO T.
XX
PI Akira S, Kishimoto T;

```
XX
DR   WPI; 1995-346089/45.
DR   N-PSDB; AAT05616.
XX
PT   New acute phase response factor - for developing inhibitory agents for
PT   treating diseases induced by cytokine(s) such as IL-6, e.g. inflammatory
PT   diseases.
XX
PS   Claim 3; Page 9-12; 31pp; English.
XX
CC   The sequence corresponds to a human acute phase response factor (APRF), a
CC   transcription factor related to signal transmission of interleukin-6 (IL-
CC   6). The protein is expressed from a human placenta cDNA, isolated using
CC   an IL-6-treated mouse liver cDNA probe. APRF-inhibitors, e.g. antibodies,
CC   antisense oligonucleotides or ribozymes, may be used to treat diseases
CC   induced by IL-6, e.g. inflammatory disease, leukemia, cancer,
CC   osteoclasia, pulmonary hypertension, etc
XX
SQ   Sequence 770 AA;

  Query Match             99.0%;  Score 792;  DB 2;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 6.5e-76;
  Matches  152;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 13
AAY03768
ID   AAY03768 standard; protein; 770 AA.
XX
AC   AAY03768;
XX
DT   11-JUN-1999  (first entry)
XX
DE   Human STAT3 allelic variant.
XX
KW   Signal Transducer and Activator of Transcription 3; STAT3; allele; IL-6;
KW   intracellular transcription factor; interleukin-6; medicament; variant;
KW   pharmaceutical; autoimmune disease; inflammatory; human.
XX
OS   Homo sapiens.
XX
PN   EP905234-A2.
XX
PD   31-MAR-1999.
XX
PF   18-FEB-1998;   98EP-00102774.
XX
PR   16-SEP-1997;   97EP-00116061.
XX
PA   (ISTF ) ARS APPLIED RES SYSTEMS HOLDING NV.
XX
PI   Serlupi-Crescenzi O,  Della Pietra L;
XX
DR   WPI; 1999-192664/17.
DR   N-PSDB; AAX29281.
XX
PT   New human Signal Transducer and Activator of Transcription 3 (STAT3)
PT   allelic variant useful for treatment of autoimmune and inflammatory
PT   disease.
XX
PS   Claim 2; Page 9-13; 32pp; English.
XX
CC   The present sequence represents a predominant allelic variant of human
CC   Signal Transducer and Activator of Transcription 3 (STAT3) protein, an
CC   intracellular transcription factor which mediates IL-6 signals. The
CC   encoding sequence differs from the original published human STAT3 gene
CC   sequence (Akira et al., 1994). Host cells comprising a recombinant STAT3
CC   DNA molecule can be used for the recombinant expression of the variant.
CC   STAT3 protein is useful as a medicament or pharmaceutical composition for
CC   treatment of autoimmune or inflammatory diseases
XX
SQ   Sequence 770 AA;

  Query Match             99.0%;  Score 792;  DB 2;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 6.5e-76;
  Matches  152;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 14
AAB19964
ID   AAB19964 standard; protein; 770 AA.
XX
AC   AAB19964;
XX
DT   28-MAR-2001  (first entry)
XX
DE   Human signal transducer and activator of transcription STAT-3.
XX
KW   STAT-3; signal transducer and activator of transcription 3; human;
KW   crystal; drug screening.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          134. .320
FT                   /note= "coiled-coil domain"
FT   Domain          321. .493
FT                   /note= "DNA-binding domain"
FT   Domain          494. .583
FT                   /note= "linker donain"
FT   Domain          584. .688
FT                   /note= "SH2 domain"
FT   Region          689. .717
FT                   /note= "C-terminal tail segment"
FT   Modified-site   705
FT                   /note= "O-phosphorylated"
XX
PN   US6160092-A.
XX
PD   12-DEC-2000.
XX
PF   29-MAY-1998;   98US-00087465.
XX
PR   29-MAY-1998;   98US-00087465.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Chen X,  Darnell JE,  Kuriyan J,  Vinkemeier U,  Zhao Y,  Jeruzalmi D;
XX
DR   WPI; 2001-101568/11.
DR   N-PSDB; AAA89229.
XX
```

PT Novel crystal useful in drug screening assays, comprises portion of
PT signal transducer, activator of transcription and duplex DNA.
XX
PS Disclosure; Col 67-71; 206pp; English.
XX
CC The present sequence is that of human signal transducer and activator of
CC transcription 3 (STAT-3). The invention provides a crystal of a core
CC portion of a STAT protein in dimer form with an 18-mer duplex DNA (see
CC AAA89233) that contains a binding site for the STAT dimer. The core
CC portion comprises a coiled-coil domain comprising 4 long helices, a DNA
CC binding domain which contains an immunoglobulin-like fold, a C-terminal
CC SH2 domain and a domain that links the DNA binding and SH2 domains. The
CC crystal is of sufficient quality to perform X-ray crystallography
CC studies. Methods of preparing the crystals are included in the invention.
CC Knowledge of the STAT protein's 3-dimensional structure will aid in
CC structure-based drug design. The crystal can be used in drug screening
CC assays to identify agonist and antagonist compounds. Antagonists can be
CC used to treat inflammation, allergy, asthma and leukaemia, and agonists
CC to treat anaemia, neutropenia, thrombocytopenia, cancer, obesity, viral
CC diseases, growth retardation, and other conditions characterized by
CC insufficient STAT activity
XX
SQ Sequence 770 AA;

Query Match 99.0%; Score 792; DB 4; Length 770;
Best Local Similarity 98.7%; Pred. No. 6.5e-76;
Matches 152; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              |||||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
```

RESULT 15
ABG69497
ID ABG69497 standard; protein; 770 AA.
XX
AC ABG69497;
XX
DT 21-OCT-2002 (first entry)
XX
DE Human bait protein STAT3.
XX
KW Human; yeast two-hybrid assay; adipocyte; bait protein; NIDDM;
KW non-insulin diabetes mellitus; obesity; selected interacting domain; SID;
KW protein-protein interaction map; PIM; anorectic; metabolic disorder.
XX
OS Homo sapiens.
XX
PN WO200253726-A2.
XX
PD 11-JUL-2002.
XX
PF 28-DEC-2001; 2001WO-EP015423.
XX
PR 02-JAN-2001; 2001US-0259377P.
XX
PA (HYBR-) HYBRIGENICS.
PA (CNRS) CENT NAT RECH SCI.
XX
PI Legrain P, Marullo S, Jockers R;
XX
DR WPI; 2002-583612/62.
DR N-PSDB; ABS51033.
XX

PT Novel complex of protein-protein interactions in adipocyte cells for
PT identifying compounds that modulate the protein-protein interactions and
PT useful for treating obesity and metabolic disorders.
XX
PS Claim 1; Page 54; 125pp; English.
XX
CC The invention relates to a complex of protein-protein interactions
CC (forming a protein-protein interaction map, PIM) in adipocyte cells as
CC defined in the specification, or polynucleotides in adipocytes encoding
CC for the polypeptides. Also included are a recombinant cell expressing the
CC interacting polypeptides and a method of selecting a modulating compound
CC in adipocyte cells, by cultivating a recombinant host cell on a selective
CC medium containing a modulating compound and a reporter gene the
CC expression of which is toxic for the recombinant host cell which is
CC transformed with two vectors, where the first vector comprises a
CC polynucleotide encoding a first hybrid polypeptide and DNA binding domain
CC and the second vector comprising a polynucleotide encoding a second
CC hybrid polypeptide and an activating domain that activates the toxic
CC reporter gene, when the first and second hybrid polypeptides interact and
CC selecting the modulating compound which inhibits the growth of the
CC recombinant host cell (i.e. using the yeast two-hybrid system). The
CC complexes are useful for identifying compounds that modulate the protein-
CC protein interactions and useful for treating obesity and metabolic
CC disorders e.g. non-insulin dependent diabetes mellitus, NIDDM. The
CC compound isolated by the method is useful for treating and preventing
CC obesity or metabolic diseases. The interactions between the proteins of
CC the complex further define a set of selected interacting domains, SID.
CC The present sequence represents a member of the protein complex of the
CC invention, used as the bait protein in the yeast two- hybrid assay
XX
SQ Sequence 770 AA;

Query Match 99.0%; Score 792; DB 5; Length 770;
Best Local Similarity 98.7%; Pred. No. 6.5e-76;
Matches 152; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              |||||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
```

Search completed: May 5, 2005, 15:01:09
Job time : 50.8306 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 5, 2005, 14:55:35 ; Search time 23.1906 Seconds
(without alignments)
495.716 Million cell updates/sec

Title: US-10-090-185-8
Perfect score: 800
Sequence: 1 MAQWNQLQQLDTRYLKQLHQ.....AVVTEKQQMLEQHLQDVRKR 154

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	800	100.0	154	3	US-09-387-418A-8	Sequence 8, Appli
2	800	100.0	770	1	US-08-369-796-12	Sequence 12, Appl
3	800	100.0	770	2	US-08-852-091-12	Sequence 12, Appl
4	800	100.0	770	2	US-08-820-754-12	Sequence 12, Appl
5	800	100.0	770	3	US-08-956-652-12	Sequence 12, Appl
6	800	100.0	770	3	US-08-956-869-12	Sequence 12, Appl
7	800	100.0	770	3	US-09-012-710-8	Sequence 8, Appli
8	800	100.0	770	3	US-08-948-547-12	Sequence 12, Appl
9	800	100.0	770	3	US-09-364-970-3	Sequence 3, Appli
10	800	100.0	770	3	US-09-364-970-5	Sequence 5, Appli
11	800	100.0	770	3	US-09-556-273-8	Sequence 8, Appli
12	800	100.0	770	3	US-08-956-653A-12	Sequence 12, Appl
13	800	100.0	770	4	US-08-212-185-12	Sequence 12, Appl
14	800	100.0	770	5	PCT-US95-17025-12	Sequence 12, Appl
15	792	99.0	770	1	US-08-416-581B-1	Sequence 1, Appli
16	792	99.0	770	1	US-08-416-581B-5	Sequence 5, Appli
17	792	99.0	770	1	US-08-416-581B-9	Sequence 9, Appli
18	792	99.0	770	3	US-09-087-465-6	Sequence 6, Appli
19	792	99.0	770	3	US-09-526-542-2	Sequence 2, Appli
20	792	99.0	770	4	US-09-972-800A-6	Sequence 6, Appli
21	792	99.0	770	4	US-10-117-087-2	Sequence 2, Appli
22	792	99.0	771	1	US-08-276-099A-14	Sequence 14, Appl
23	792	99.0	771	1	US-08-781-890-14	Sequence 14, Appl
24	380	47.5	154	3	US-09-387-418A-11	Sequence 11, Appl
25	380	47.5	712	1	US-08-369-796-6	Sequence 6, Appli
26	380	47.5	712	2	US-08-852-091-6	Sequence 6, Appli
27	380	47.5	712	2	US-08-820-754-6	Sequence 6, Appli
28	380	47.5	712	3	US-08-956-652-6	Sequence 6, Appli
29	380	47.5	712	3	US-08-956-869-6	Sequence 6, Appli
30	380	47.5	712	3	US-08-948-547-6	Sequence 6, Appli
31	380	47.5	712	3	US-08-956-653A-6	Sequence 6, Appli
32	380	47.5	712	4	US-08-212-185-6	Sequence 6, Appli
33	380	47.5	712	4	US-09-430-806A-2	Sequence 2, Appli
34	380	47.5	712	5	PCT-US95-17025-6	Sequence 6, Appli
35	380	47.5	729	4	US-09-917-254-97	Sequence 97, Appl
36	380	47.5	740	1	US-08-276-099A-12	Sequence 12, Appl
37	380	47.5	740	1	US-08-781-890-12	Sequence 12, Appl
38	380	47.5	750	1	US-08-369-796-4	Sequence 4, Appli
39	380	47.5	750	2	US-08-852-091-4	Sequence 4, Appli
40	380	47.5	750	2	US-08-820-754-4	Sequence 4, Appli
41	380	47.5	750	3	US-08-956-652-4	Sequence 4, Appli
42	380	47.5	750	3	US-08-956-869-4	Sequence 4, Appli
43	380	47.5	750	3	US-08-948-547-4	Sequence 4, Appli
44	380	47.5	750	3	US-09-087-465-2	Sequence 2, Appli
45	380	47.5	750	3	US-09-364-970-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-387-418A-8
; Sequence 8, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/09/387,418A
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-387-418A-8

Query Match 100.0%; Score 800; DB 3; Length 154;
Best Local Similarity 100.0%; Pred. No. 7.3e-79;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
```

RESULT 2
US-08-369-796-12
; Sequence 12, Application US/08369796
; Patent No. 5716622
; GENERAL INFORMATION:
; APPLICANT: James E. Darnell, Jr.
; APPLICANT: Zilong Wen
; APPLICANT: Curt M. Horvath
; APPLICANT: Zhong Zhong
; TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
; TITLE OF INVENTION: TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS

```
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/369,796
;      FILING DATE:  06-JAN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-116
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-369-796-12

  Query Match             100.0%;  Score 800;  DB 1;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-78;
  Matches  154;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 3
US-08-852-091-12
; Sequence 12, Application US/08852091
; Patent No. 5883228
;  GENERAL INFORMATION:
;    APPLICANT:  James E. Darnell, Jr.
;    APPLICANT:  Zilong Wen
;    APPLICANT:  Curt M. Horvath
;    APPLICANT:  Zhong Zhong
;    TITLE OF INVENTION:  FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
;    TITLE OF INVENTION:  TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/852,091
;      FILING DATE:  06-MAY-1997
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/369,796
;      FILING DATE:  06-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-116
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-852-091-12

  Query Match             100.0%;  Score 800;  DB 2;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-78;
  Matches  154;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 4
US-08-820-754-12
; Sequence 12, Application US/08820754
; Patent No. 5976835
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/820,754
;      FILING DATE:  19-MAR-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,185
;      FILING DATE:  11-MAR-1994
```

```
;       APPLICATION NUMBER:  US 07/980,498
;       FILING DATE:  23-NOV-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/854,296
;       FILING DATE:  19-MAR-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  WO US93/02569
;       FILING DATE:  19-MAR-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/126,588
;       FILING DATE:  24-SEP-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Jackson Esq., David A.
;       REGISTRATION NUMBER:  26,742
;       REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  201 487-5800
;       TELEFAX:  201 343-1684
;       TELEX:  133521
;   INFORMATION FOR SEQ ID NO:  12:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  770 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-820-754-12

  Query Match             100.0%;  Score 800;  DB 2;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-78;
  Matches  154;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 5
US-08-956-652-12
; Sequence 12, Application US/08956652
; Patent No. 6013475
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/956,652
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,185
;      FILING DATE:  11-MAR-1994
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-956-652-12

  Query Match             100.0%;  Score 800;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-78;
  Matches  154;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 6
US-08-956-869-12
; Sequence 12, Application US/08956869
; Patent No. 6030808
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  US/08/956,869
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/212,185
;       FILING DATE:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/854,296
;       FILING DATE:  19-MAR-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  WO US93/02569
;       FILING DATE:  19-MAR-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/126,588
;       FILING DATE:  24-SEP-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Jackson Esq., David A.
;       REGISTRATION NUMBER:  26,742
;       REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  201 487-5800
;       TELEFAX:  201 343-1684
;       TELEX:  133521
;   INFORMATION FOR SEQ ID NO:  12:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  770 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-956-869-12

  Query Match             100.0%;  Score 800;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-78;
  Matches  154;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 7
US-09-012-710-8
; Sequence 8, Application US/09012710
; Patent No. 6087478
;  GENERAL INFORMATION:
;    APPLICANT:  Vinkemeier, Uwe
;    APPLICANT:  Moarefi, Ismail
;    APPLICANT:  Darnell, Jr., James E.
;    APPLICANT:  Kuriyan, John
;    TITLE OF INVENTION:  A CRYSTAL OF THE N-TERMINAL DOMAIN OF A
;    TITLE OF INVENTION:  STAT PROTEIN AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/012,710
;       FILING DATE:
;       CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;       NAME:  Jackson Esq., David A.
;       REGISTRATION NUMBER:  26,742
;       REFERENCE/DOCKET NUMBER:  600-1-194
;    TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  201-487-5800
;       TELEFAX:  201-343-1684
;       TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH:  770 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-09-012-710-8

  Query Match             100.0%;  Score 800;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-78;
  Matches  154;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 8
US-08-948-547-12
; Sequence 12, Application US/08948547
; Patent No. 6124118
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/948,547
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/212,185
;      FILING DATE:  11-MAR-1994
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-948-547-12

  Query Match             100.0%;  Score 800;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-78;
  Matches  154;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              |||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 9
US-09-364-970-3
; Sequence 3, Application US/09364970
; Patent No. 6235873
; GENERAL INFORMATION:
;  APPLICANT: Bromberg, Jacqueline
;  TITLE OF INVENTION: CONSTITUTIVELY ACTIVE TRANSCRIPTION FACTORS AND THEIR
;  TITLE OF INVENTION: USES FOR IDENTIFYING MODULATORS OF ACTIVITY INCLUDING
;  TITLE OF INVENTION: DYSPROLIFERATIVE CELLULAR CHANGES
;  FILE REFERENCE: 600-1-252
;  CURRENT APPLICATION NUMBER: US/09/364,970
;  CURRENT FILING DATE: 1999-07-31
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-364-970-3

  Query Match             100.0%;  Score 800;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-78;
  Matches  154;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              |||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 10
US-09-364-970-5
; Sequence 5, Application US/09364970
; Patent No. 6235873
; GENERAL INFORMATION:
;  APPLICANT: Bromberg, Jacqueline
;  TITLE OF INVENTION: CONSTITUTIVELY ACTIVE TRANSCRIPTION FACTORS AND THEIR
;  TITLE OF INVENTION: USES FOR IDENTIFYING MODULATORS OF ACTIVITY INCLUDING
;  TITLE OF INVENTION: DYSPROLIFERATIVE CELLULAR CHANGES
;  FILE REFERENCE: 600-1-252
;  CURRENT APPLICATION NUMBER: US/09/364,970
;  CURRENT FILING DATE: 1999-07-31
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-364-970-5

  Query Match             100.0%;  Score 800;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-78;
  Matches  154;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              |||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 11
US-09-556-273-8
; Sequence 8, Application US/09556273
; Patent No. 6312887
;  GENERAL INFORMATION:
;    APPLICANT:  Vinkemeier, Uwe
;    APPLICANT:  Moarefi, Ismail
;    APPLICANT:  Darnell, Jr., James E.
;    APPLICANT:  Kuriyan, John
;    TITLE OF INVENTION:  A CRYSTAL OF THE N-TERMINAL DOMAIN OF A
;    TITLE OF INVENTION:  STAT PROTEIN AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/556,273
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/012,710
;      FILING DATE:
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-194
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-09-556-273-8

  Query Match            100.0%;  Score 800;  DB 3;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 5.9e-78;
  Matches  154;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 12
US-08-956-653A-12
; Sequence 12, Application US/08956653A
; Patent No. 6338949
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/956,653A
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,185
;      FILING DATE:  11-MAR-1994
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-195
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-956-653A-12

  Query Match            100.0%;  Score 800;  DB 3;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 5.9e-78;
  Matches  154;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 13
US-08-212-185-12
; Sequence 12, Application US/08212185
; Patent No. 6605442
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/212,185
;      FILING DATE:  11-MAR-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-212-185-12

  Query Match             100.0%;  Score 800;  DB 4;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-78;
  Matches  154;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 14
PCT-US95-17025-12
; Sequence 12, Application PC/TUS9517025
;  GENERAL INFORMATION:
;    APPLICANT:  James E. Darnell, Jr.
;    APPLICANT:  Zilong Wen
;    APPLICANT:  Curt M. Horvath
;    APPLICANT:  Zhong Zhong
;    TITLE OF INVENTION:  FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
;    TITLE OF INVENTION:  TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/17025
;      FILING DATE:  28-DEC-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/369,796
;      FILING DATE:  06-JAN-1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-116
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US95-17025-12

  Query Match             100.0%;  Score 800;  DB 5;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 5.9e-78;
  Matches  154;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 15
US-08-416-581B-1
; Sequence 1, Application US/08416581B
; Patent No. 5719042
;  GENERAL INFORMATION:
;    APPLICANT:  Kishimoto, Tadamitsu
;    APPLICANT:  Akira, Shizuo
;    TITLE OF INVENTION:  TRANSCRIPTION FACTOR APRF
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SUGHRUE, MION, ZINN, MACPEAK & SEAS
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/416,581B
;      FILING DATE:  04-APR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-65825/1994
;      FILING DATE:  04-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Nakamura, Dean H.
;      REGISTRATION NUMBER:  33,981
;      REFERENCE/DOCKET NUMBER:  Q-37891
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)293-7060
;      TELEFAX:  (202)293-7860
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-416-581B-1
```

Query Match 99.0%; Score 792; DB 1; Length 770;
Best Local Similarity 98.7%; Pred. No. 4.3e-77;
Matches 152; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
```

Search completed: May 5, 2005, 14:58:51
Job time : 26.1906 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2005, 14:59:00 ; Search time 64.8612 Seconds
(without alignments)
792.064 Million cell updates/sec

Title: US-10-090-185-8
Perfect score: 800
Sequence: 1 MAQWNQLQQLDTRYLKQLHQ.....AVVTEKQQMLEQHLQDVRKR 154

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1428581 seqs, 333598853 residues

Total number of hits satisfying chosen parameters: 1428581

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	800	100.0	154	13	US-10-090-185-8	Sequence 8, Appli
2	800	100.0	770	11	US-09-876-773-12	Sequence 12, Appl
3	800	100.0	770	14	US-10-045-792-8	Sequence 8, Appli
4	800	100.0	770	17	US-10-639-617-12	Sequence 12, Appl
5	792	99.0	720	15	US-10-380-020-4	Sequence 4, Appli
6	792	99.0	769	15	US-10-380-020-2	Sequence 2, Appli
7	792	99.0	769	15	US-10-380-020-5	Sequence 5, Appli
8	792	99.0	770	14	US-10-038-010-56	Sequence 56, Appl
9	792	99.0	770	14	US-10-117-087-2	Sequence 2, Appli
10	792	99.0	770	15	US-10-116-275-329	Sequence 329, App
11	792	99.0	770	15	US-10-116-275-349	Sequence 349, App
12	792	99.0	793	9	US-09-925-302-780	Sequence 780, App
13	792	99.0	793	10	US-09-925-302-780	Sequence 780, App
14	380	47.5	154	13	US-10-090-185-11	Sequence 11, Appl
15	380	47.5	712	11	US-09-876-773-6	Sequence 6, Appli
16	380	47.5	712	14	US-10-245-120-2	Sequence 2, Appli
17	380	47.5	712	17	US-10-639-617-6	Sequence 6, Appli
18	380	47.5	712	17	US-10-936-390-5	Sequence 5, Appli
19	380	47.5	750	9	US-09-833-205-2	Sequence 2, Appli
20	380	47.5	750	11	US-09-876-773-4	Sequence 4, Appli
21	380	47.5	750	14	US-10-245-120-1	Sequence 1, Appli
22	380	47.5	750	14	US-10-308-279-44	Sequence 44, Appl
23	380	47.5	750	16	US-10-755-889-352	Sequence 352, App
24	380	47.5	750	16	US-10-755-889-823	Sequence 823, App
25	380	47.5	750	17	US-10-492-043-19	Sequence 19, Appl
26	380	47.5	750	17	US-10-639-617-4	Sequence 4, Appli
27	380	47.5	786	9	US-09-925-297-550	Sequence 550, App
28	378	47.2	749	9	US-09-833-205-4	Sequence 4, Appli
29	378	47.2	749	9	US-09-833-205-6	Sequence 6, Appli
30	378	47.2	749	11	US-09-876-773-8	Sequence 8, Appli
31	378	47.2	749	14	US-10-045-792-7	Sequence 7, Appli
32	378	47.2	749	14	US-10-205-194-39	Sequence 39, Appl
33	378	47.2	749	17	US-10-639-617-8	Sequence 8, Appli
34	360	45.0	131	14	US-10-245-120-4	Sequence 4, Appli
35	334.5	41.8	748	11	US-09-876-773-10	Sequence 10, Appl
36	334.5	41.8	748	17	US-10-639-617-10	Sequence 10, Appl
37	334.5	41.8	749	14	US-10-045-792-9	Sequence 9, Appli
38	333.5	41.7	423	14	US-10-177-293-437	Sequence 437, App
39	333.5	41.7	748	14	US-10-177-293-439	Sequence 439, App
40	333.5	41.7	748	16	US-10-755-889-70	Sequence 70, Appl
41	333.5	41.7	748	16	US-10-755-889-362	Sequence 362, App
42	247.5	30.9	532	14	US-10-205-823-389	Sequence 389, App
43	247.5	30.9	851	11	US-09-876-773-2	Sequence 2, Appli
44	247.5	30.9	851	14	US-10-045-792-2	Sequence 2, Appli
45	247.5	30.9	851	15	US-10-179-451-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-10-090-185-8
; Sequence 8, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR F
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/10/090,185
; CURRENT FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: 09/387,418
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-090-185-8

Query Match 100.0%; Score 800; DB 13; Length 154;
Best Local Similarity 100.0%; Pred. No. 4.7e-77;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
```

```
Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 2
US-09-876-773-12
; Sequence 12, Application US/09876773
; Publication No. US20040058318A1
;   GENERAL INFORMATION:
;        APPLICANT: Darnell Jr., James E.
;                   Schindler, Christian W.
;                   Fu, Xian-Yuan
;                   Wen, Zilong
;                   Zhong, Zhong
;        TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
;                            SEQUENCES AND METHODS OF USE THEREOF
;        NUMBER OF SEQUENCES: 25
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Klauber & Jackson
;             STREET: 411 Hackensack Avenue
;             CITY: Hackensack
;             STATE: New Jersey
;             COUNTRY: USA
;             ZIP: 07601
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/876,773
;             FILING DATE: 07-Jun-2001
;             CLASSIFICATION: 435
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/212,185
;             FILING DATE: 11-MAR-1994
;             APPLICATION NUMBER: US 07/980,498
;             FILING DATE: 23-NOV-1992
;             APPLICATION NUMBER: US 07/854,296
;             FILING DATE: 19-MAR-1992
;             APPLICATION NUMBER: WO US93/02569
;             FILING DATE: 19-MAR-1993
;             APPLICATION NUMBER: US 08/126,588
;             FILING DATE: 24-SEP-1993
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Jackson Esq., David A.
;             REGISTRATION NUMBER: 26,742
;             REFERENCE/DOCKET NUMBER: 600-1-073 CIP
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 201 487-5800
;             TELEFAX: 201 343-1684
;             TELEX: 133521
;    INFORMATION FOR SEQ ID NO: 12:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 770 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-876-773-12

  Query Match             100.0%;  Score 800;  DB 11;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 3.7e-76;
  Matches  154;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 3
US-10-045-792-8
; Sequence 8, Application US/10045792
; Publication No. US20030003563A1
;   GENERAL INFORMATION:
;        APPLICANT: Vinkemeier, Uwe
;                   Moarefi, Ismail
;                   Darnell, Jr., James E.
;                   Kuriyan, John
;        TITLE OF INVENTION: A CRYSTAL OF THE N-TERMINAL DOMAIN OF A
;                            STAT PROTEIN AND METHODS OF USE THEREOF
;        NUMBER OF SEQUENCES: 13
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Klauber & Jackson
;             STREET: 411 Hackensack Avenue, 4th Floor
;             CITY: Hackensack
;             STATE: New Jersey
;             COUNTRY: USA
;             ZIP: 07601
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/045,792
;             FILING DATE: 19-Oct-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/012,710
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Jackson Esq., David A.
;             REGISTRATION NUMBER: 26,742
;             REFERENCE/DOCKET NUMBER: 600-1-194
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 201-487-5800
;             TELEFAX: 201-343-1684
;             TELEX: 133521
;    INFORMATION FOR SEQ ID NO: 8:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 770 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        HYPOTHETICAL: NO
;        SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-045-792-8

  Query Match             100.0%;  Score 800;  DB 14;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 3.7e-76;
  Matches  154;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
```

```
RESULT 4
US-10-639-617-12
; Sequence 12, Application US/10639617
; Publication No. US20050079543A1
;   GENERAL INFORMATION:
;        APPLICANT: Darnell Jr., James E.
;                   Schindler, Christian W.
;                   Fu, Xian-Yuan
;                   Wen, Zilong
;                   Zhong, Zhong
;        TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
;                            SEQUENCES AND METHODS OF USE THEREOF
;        NUMBER OF SEQUENCES: 25
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Klauber & Jackson
;             STREET: 411 Hackensack Avenue
;             CITY: Hackensack
;             STATE: New Jersey
;             COUNTRY: USA
;             ZIP: 07601
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/639,617
;             FILING DATE: 12-Aug-2003
;             CLASSIFICATION: 435
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/212,185
;             FILING DATE: 11-MAR-1994
;             APPLICATION NUMBER: US 07/980,498
;             FILING DATE: 23-NOV-1992
;             APPLICATION NUMBER: US 07/854,296
;             FILING DATE: 19-MAR-1992
;             APPLICATION NUMBER: WO US93/02569
;             FILING DATE: 19-MAR-1993
;             APPLICATION NUMBER: US 08/126,588
;             FILING DATE: 24-SEP-1993
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Jackson Esq., David A.
;             REGISTRATION NUMBER: 26,742
;             REFERENCE/DOCKET NUMBER: 600-1-073 CIP
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 201 487-5800
;             TELEFAX: 201 343-1684
;             TELEX: 133521
;   INFORMATION FOR SEQ ID NO: 12:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 770 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-639-617-12

  Query Match            100.0%;  Score 800;  DB 17;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 3.7e-76;
  Matches  154;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
```

```
RESULT 5
US-10-380-020-4
; Sequence 4, Application US/10380020
; Publication No. US20040052762A1
; GENERAL INFORMATION:
;  APPLICANT: Yu, Hua
;  APPLICANT:  Pardoll, Drew
;  APPLICANT:  Jove, Richard
;  APPLICANT:  Dalton, Willian
;  TITLE OF INVENTION: Stat3 Agonists and Antagonists and Therapeutic Uses Thereof
;  FILE REFERENCE: 10873-009-999
;  CURRENT APPLICATION NUMBER: US/10/380,020
;  CURRENT FILING DATE:  2003-03-07
;  PRIOR APPLICATION NUMBER: 60/231,212
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
;   LENGTH: 720
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-380-020-4

  Query Match            99.0%;  Score 792;  DB 15;  Length 720;
  Best Local Similarity  98.7%;  Pred. No. 2.4e-75;
  Matches  152;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
```

```
RESULT 6
US-10-380-020-2
; Sequence 2, Application US/10380020
; Publication No. US20040052762A1
; GENERAL INFORMATION:
;  APPLICANT: Yu, Hua
;  APPLICANT:  Pardoll, Drew
;  APPLICANT:  Jove, Richard
;  APPLICANT:  Dalton, Willian
;  TITLE OF INVENTION: Stat3 Agonists and Antagonists and Therapeutic Uses Thereof
;  FILE REFERENCE: 10873-009-999
;  CURRENT APPLICATION NUMBER: US/10/380,020
;  CURRENT FILING DATE:  2003-03-07
;  PRIOR APPLICATION NUMBER: 60/231,212
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 769
;   TYPE: PRT<
;   ORGANISM: Homo Sapiens
US-10-380-020-2

  Query Match            99.0%;  Score 792;  DB 15;  Length 769;
  Best Local Similarity  98.7%;  Pred. No. 2.7e-75;
  Matches  152;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy         121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
                |||||||||||||||||||||||||||||||||||
Db         121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 7
US-10-380-020-5
; Sequence 5, Application US/10380020
; Publication No. US20040052762A1
; GENERAL INFORMATION:
;  APPLICANT: Yu, Hua
;  APPLICANT:  Pardoll, Drew
;  APPLICANT:  Jove, Richard
;  APPLICANT:  Dalton, Willian
;  TITLE OF INVENTION: Stat3 Agonists and Antagonists and Therapeutic Uses Thereof
;  FILE REFERENCE: 10873-009-999
;  CURRENT APPLICATION NUMBER: US/10/380,020
;  CURRENT FILING DATE:  2003-03-07
;  PRIOR APPLICATION NUMBER: 60/231,212
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 769
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-380-020-5

  Query Match             99.0%;  Score 792;  DB 15;  Length 769;
  Best Local Similarity   98.7%;  Pred. No. 2.7e-75;
  Matches  152;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              |||||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              |||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 8
US-10-038-010-56
; Sequence 56, Application US/10038010
; Publication No. US20030040089A1
; GENERAL INFORMATION:
;  APPLICANT: HYBRIGENICS
;  APPLICANT:  Pierre, Legrain
;  TITLE OF INVENTION: Protein-protein interactions in adipocyte cells
;  FILE REFERENCE: B4767A
;  CURRENT APPLICATION NUMBER: US/10/038,010
;  CURRENT FILING DATE:  2002-07-23
;  PRIOR APPLICATION NUMBER: US 60/259,377
;  PRIOR FILING DATE: 2001-01-02
;  NUMBER OF SEQ ID NOS: 67
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: STAT3 : Transcription factor
;   LOCATION: (1)..(770)
;   OTHER INFORMATION:
US-10-038-010-56

  Query Match             99.0%;  Score 792;  DB 14;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 2.7e-75;
  Matches  152;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              |||||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              |||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 9
US-10-117-087-2
; Sequence 2, Application US/10117087
; Publication No. US20030166854A1
; GENERAL INFORMATION:
;  APPLICANT: SERLUPI-CRESCENZI, Ottaviano
;  APPLICANT:  DELLA PIETRA, Linda
;  TITLE OF INVENTION: ALLELIC VARIANTS OF HUMAN STAT3
;  FILE REFERENCE: SERLUPI=2
;  CURRENT APPLICATION NUMBER: US/10/117,087
;  CURRENT FILING DATE:  2002-04-08
;  PRIOR APPLICATION NUMBER: US/09/526,542
;  PRIOR FILING DATE: 2000-03-19
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Human
US-10-117-087-2

  Query Match             99.0%;  Score 792;  DB 14;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 2.7e-75;
  Matches  152;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              |||||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              |||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 10
US-10-116-275-329
; Sequence 329, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
;  APPLICANT: Elan Pharmaceutical Technology
;  APPLICANT:  O'Mahony, Daniel J.
;  APPLICANT:  Brayden, David
;  APPLICANT:  Byrne, Daragh
;  APPLICANT:  Lambkin, Imelda
;  APPLICANT:  Higgins, Lisa
;  TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
;  TITLE OF INVENTION:  Compositions Targeting Peyer's Patches and M Cell Receptors
;  FILE REFERENCE: E1067/20087
;  CURRENT APPLICATION NUMBER: US/10/116,275
;  CURRENT FILING DATE:  2002-10-04
;  NUMBER OF SEQ ID NOS: 349
;  SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 329
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Rat
US-10-116-275-329

  Query Match             99.0%;  Score 792;  DB 15;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 2.7e-75;
  Matches  152;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              |||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 11
US-10-116-275-349
; Sequence 349, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
;  APPLICANT: Elan Pharmaceutical Technology
;  APPLICANT:  O'Mahony, Daniel J.
;  APPLICANT:  Brayden, David
;  APPLICANT:  Byrne, Daragh
;  APPLICANT:  Lambkin, Imelda
;  APPLICANT:  Higgins, Lisa
;  TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
;  TITLE OF INVENTION:  Compositions Targeting Peyer's Patches and M Cell Receptors
;  FILE REFERENCE: E1067/20087
;  CURRENT APPLICATION NUMBER: US/10/116,275
;  CURRENT FILING DATE:  2002-10-04
;  NUMBER OF SEQ ID NOS: 349
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 349
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-116-275-349

  Query Match             99.0%;  Score 792;  DB 15;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 2.7e-75;
  Matches  152;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              |||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 12
US-09-925-302-780
; Sequence 780, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA104
;  CURRENT APPLICATION NUMBER: US/09/925,302
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05918
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 896
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 780
;   LENGTH: 793
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-925-302-780

  Query Match             99.0%;  Score 792;  DB 9;  Length 793;
  Best Local Similarity   98.7%;  Pred. No. 2.8e-75;
  Matches  152;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db         24 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 83

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 143

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              |||||||||||||||||||||||||||||||||
Db        144 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 177


RESULT 13
US-09-925-302-780
; Sequence 780, Application US/09925302
; Publication No. US20030064072A9
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA104
;  CURRENT APPLICATION NUMBER: US/09/925,302
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05918
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 896
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 780
;   LENGTH: 793
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-925-302-780

  Query Match             99.0%;  Score 792;  DB 10;  Length 793;
  Best Local Similarity   98.7%;  Pred. No. 2.8e-75;
  Matches  152;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db         24 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 83

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 143

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              |||||||||||||||||||||||||||||||||
Db        144 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 177


RESULT 14
US-10-090-185-11
; Sequence 11, Application US/10090185
; Publication No. US20020197647A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 154
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-11

  Query Match             47.5%;  Score 380;  DB 13;  Length 154;
  Best Local Similarity   44.8%;  Pred. No. 2.5e-32;
  Matches   69;  Conservative   42;  Mismatches   41;  Indels    2;  Gaps    1;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              |:|| :|||||:::|:|:|||| |:||||:||:|| |:| ||| :||:  | ||: ||:|
Db          1 MSQWYELQQLDSKFLEQVHQLYDDSFPMEIRQYLAQWLEKQDWEHAANDVSFATIRFHDL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              | ::| |||||  |:| | |||:|: |: ||  : | |:::: |:  || || ::|: |
Db         61 LSQLDDQYSRFSLENNFLLQHNIRKSKRNLQDNFQEDPIQMSMIIYSCLKEERKILENAQ 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
                |  |: :  : |: :||: |:  :::|: :
Db        121 RFNQ--AQSGNIQSTVMLDKQKELDSKVRNVKDK 152


RESULT 15
US-09-876-773-6
; Sequence 6, Application US/09876773
; Publication No. US20040058318A1
;  GENERAL INFORMATION:
;       APPLICANT: Darnell Jr., James E.
;                  Schindler, Christian W.
;                  Fu, Xian-Yuan
;                  Wen, Zilong
;                  Zhong, Zhong
;       TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
;                           SEQUENCES AND METHODS OF USE THEREOF
;       NUMBER OF SEQUENCES: 25
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Klauber & Jackson
;            STREET: 411 Hackensack Avenue
;            CITY: Hackensack
;            STATE: New Jersey
;            COUNTRY: USA
;            ZIP: 07601
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/876,773
;            FILING DATE: 07-Jun-2001
;            CLASSIFICATION: 435
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/212,185
;            FILING DATE: 11-MAR-1994
;            APPLICATION NUMBER: US 07/980,498
;            FILING DATE: 23-NOV-1992
;            APPLICATION NUMBER: US 07/854,296
;            FILING DATE: 19-MAR-1992
;            APPLICATION NUMBER: WO US93/02569
;            FILING DATE: 19-MAR-1993
;            APPLICATION NUMBER: US 08/126,588
;            FILING DATE: 24-SEP-1993
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Jackson Esq., David A.
;            REGISTRATION NUMBER: 26,742
;            REFERENCE/DOCKET NUMBER: 600-1-073 CIP
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 201 487-5800
;            TELEFAX: 201 343-1684
;            TELEX: 133521
;  INFORMATION FOR SEQ ID NO: 6:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 712 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-876-773-6

  Query Match             47.5%;  Score 380;  DB 11;  Length 712;
  Best Local Similarity   44.8%;  Pred. No. 1.8e-31;
  Matches   69;  Conservative   42;  Mismatches   41;  Indels    2;  Gaps    1;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              |:|| :|||||:::|:|:|||| |:||||:||:|| |:| ||| :||:  | ||: ||:|
Db          1 MSQWYELQQLDSKFLEQVHQLYDDSFPMEIRQYLAQWLEKQDWEHAANDVSFATIRFHDL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              | ::| |||||  |:| | |||:|: |: ||  : | |:::: |:  || || ::|: |
Db         61 LSQLDDQYSRFSLENNFLLQHNIRKSKRNLQDNFQEDPIQMSMIIYSCLKEERKILENAQ 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
                |  |: :  : |: :||: |:  :::|: :
Db        121 RFNQ--AQSGNIQSTVMLDKQKELDSKVRNVKDK 152


Search completed: May  5, 2005, 15:17:40
Job time : 65.8612 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 5, 2005, 14:55:35 ; Search time 11.2329 Seconds
(without alignments)
1319.101 Million cell updates/sec

Title: US-10-090-185-8
Perfect score: 800
Sequence: 1 MAQWNQLQQLDTRYLKQLHQ.....AVVTEKQQMLEQHLQDVRKR 154

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	792	99.0	770	2	A54444	DNA-binding protei
2	792	99.0	770	2	I49508	ISGF3 p91-related
3	380	47.5	739	2	A46159	interferon-depende
4	334.5	41.8	748	2	A56047	gamma-interferon a
5	247.5	30.9	851	2	A46160	interferon alpha-i
6	188	23.5	786	2	I49274	mammary gland fact
7	187	23.4	794	2	G02317	transcription acti
8	186	23.2	793	2	S54772	mammary gland fact
9	147.5	18.4	794	2	S55527	mammary gland fact
10	98	12.2	331	2	AI0354	probable lipoprote
11	97	12.1	837	2	I57557	DNA-Binding Protei
12	91.5	11.4	848	2	A54740	interleukin-4-indu
13	88	11.0	1355	2	T22552	hypothetical prote
14	82.5	10.3	371	2	T40287	hypothetical prote
15	82.5	10.3	373	2	F71884	probable lipopolys
16	82	10.2	498	2	C69587	L-arabinose isomer
17	82	10.2	620	2	S56790	probable membrane
18	82	10.2	764	2	C83513	hypothetical prote
19	81.5	10.2	276	2	H81707	inclusion membrane
20	81	10.1	2954	2	T14156	kinesin-related pr
21	80.5	10.1	1110	2	A43253	Large tra-1 protei
22	80	10.0	508	1	KRSHL2	keratin type II, m
23	79.5	9.9	392	2	T19867	hypothetical prote
24	79.5	9.9	444	2	E90192	DNA repair protein
25	79.5	9.9	613	2	S48557	hypothetical prote
26	79.5	9.9	638	2	I53169	cytokeratin 2 - hu
27	79.5	9.9	4957	2	T03455	ALR protein - huma
28	79.5	9.9	5262	2	T03454	ALR protein - huma
29	79	9.9	4540	2	T30838	cytoplasmic dynein
30	78.5	9.8	1033	2	S73693	MG328 homolog P01_
31	78.5	9.8	2101	2	A42184	nuclear mitotic ap
32	78.5	9.8	2168	2	T30171	ninein - mouse
33	78	9.8	409	2	F71956	hypothetical prote
34	78	9.8	629	2	A29666	keratin, 65K type
35	78	9.8	1462	2	T00345	hypothetical prote
36	77.5	9.7	2133	2	T30637	hypothetical prote
37	77	9.6	191	2	AB2247	hypothetical prote
38	77	9.6	2160	2	T20241	hypothetical prote
39	76.5	9.6	625	2	T37604	probable transcrip
40	76	9.5	390	2	A72108	hypothetical prote
41	76	9.5	390	2	B86514	hypothetical prote
42	76	9.5	442	2	C83192	conserved hypothet
43	76	9.5	812	2	T34180	hypothetical prote
44	75.5	9.4	895	2	T13414	probable alpha-act
45	75.5	9.4	924	2	T13413	probable alpha-act

ALIGNMENTS

RESULT 1
A54444
DNA-binding protein APRF - human
C;Species: Homo sapiens (man)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 09-Jul-2004
C;Accession: A54444
R;Akira, S.; Nishio, Y.; Inoue, M.; Wang, X.J.; Wei, S.; Matsusaka, T.; Yoshida, K.; Sud
Cell 77, 63-71, 1994
A;Title: Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related tra
A;Reference number: A54444; MUID:94208062; PMID:7512451
A;Accession: A54444
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-770 <RES>
A;Cross-references: UNIPROT:P40763; GB:L29277; NID:g475788; PID:g475789
C;Genetics:
A;Gene: GDB:STAT3; APRF
A;Cross-references: GDB:358950
A;Map position: 17q21-17q21
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: DNA binding; transcription factor

Query Match 99.0%; Score 792; DB 2; Length 770;
Best Local Similarity 98.7%; Pred. No. 1.1e-61;
Matches 152; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
```

RESULT 2
I49508
ISGF3 p91-related transcription factor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I49508; I49009
R;Akira, S.; Nishio, Y.; Inoue, M.; Wang, X.J.; Wei, S.; Matsusaka, T.; Yoshida, K.; Sud
Cell 77, 63-71, 1994
A;Title: Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related tra
A;Reference number: A54444; MUID:94208062; PMID:7512451
A;Accession: I49508
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA

```
A;Residues: 1-770 <RES>
A;Cross-references: UNIPROT:P42227; GB:L29278; NID:g476715; PIDN:AAA37254.1; PID:g476716
R;Raz, R.; Durbin, J.E.; Levy, D.E.
J. Biol. Chem. 269, 24391-24395, 1994
A;Title: Acute phase response factor and additional members of the interferon-stimulated
A;Reference number: I49009; MUID:95014185; PMID:7523373
A;Accession: I49009
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-393,'M',395-700,702-770 <RE2>
A;Cross-references: EMBL:U08378; NID:g473889; PIDN:AAA56668.1; PID:g473890
C;Genetics:
A;Gene: APRF
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             99.0%;  Score 792;  DB 2;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 1.1e-61;
  Matches  152;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 3
A46159
interferon-dependent positive-acting transcription factor ISGF-3 91K chain - human
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A46159
R;Schindler, C.; Fu, X.Y.; Improta, T.; Aebersold, R.; Darnell Jr., J.E.
Proc. Natl. Acad. Sci. U.S.A. 89, 7836-7839, 1992
A;Title: Proteins of transcription factor ISGF-3: one gene encodes the 91-and 84-kDa ISG
A;Reference number: A46159; MUID:92366557; PMID:1502203
A;Accession: A46159
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid; protein
A;Residues: 1-739 <SCH>
A;Cross-references: UNIPROT:P42224
A;Experimental source: HeLa cells
A;Note: sequence extracted from NCBI backbone (NCBIP:110818)
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             47.5%;  Score 380;  DB 2;  Length 739;
  Best Local Similarity   44.8%;  Pred. No. 1.5e-25;
  Matches   69;  Conservative   42;  Mismatches   41;  Indels    2;  Gaps    1;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              |:|| :|||||:::|:|:|||| |:||||:||:|| |:| ||| :||:  | ||: ||:|
Db          1 MSQWYELQQLDSKFLEQVHQLYDDSFPMEIRQYLAQWLEKQDWEHAANDVSFATIRFHDL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              | ::| |||||  |:| | |||:|: |: ||  : | |:::: |:  || || ::|: |
Db         61 LSQLDDQYSRFSLENNFLLQHNIRKSKRNLQDNFQEDPIQMSMIIYSCLKEERKILENAQ 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
                 |   |: :  : |: :||: |:  :::|: :
Db        121 RFNQ--AQSGNIQSTVMLDKQKELDSKVRNVKDK 152


RESULT 4
A56047
gamma-interferon activation site-binding protein Stat4 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004
C;Accession: A56047
R;Yamamoto, K.; Quelle, F.W.; Thierfelder, W.E.; Kreider, B.L.; Gilbert, D.J.; Jenkins, N
Mol. Cell. Biol. 14, 4342-4349, 1994
A;Title: Stat4, a novel gamma interferon activation site-binding protein expressed in ear
A;Reference number: A56047; MUID:94277038; PMID:8007943
A;Accession: A56047
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-748 <YAM>
A;Cross-references: UNIPROT:P42228; GB:U09351; NID:g509502; PIDN:AAA19692.1; PID:g509503
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: DNA binding; phosphoprotein

  Query Match             41.8%;  Score 334.5;  DB 2;  Length 748;
  Best Local Similarity   41.4%;  Pred. No. 1.5e-21;
  Matches   63;  Conservative   38;  Mismatches   50;  Indels    1;  Gaps    1;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              |:||||:|||: ::|:|:  | | | ||||:|  || |||:|||  |:: |: ||::  ||
Db          1 MSQWNQVQQLEIKFLEQVDQFYDDNFPMEIRHLLAQWIETQDWEVASNNETMATILLQNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              | ::|:|  |  :| |:|  |||:||:: || ::   || :| ::: || || |:|  ||
Db         61 LIQLDEQLGRVSKEKNLLLIHNLKRIKRVLQGKFHGNPMHVAVVISNCLREERRIL-AAA 119

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVR 152
                 ||       :: |:|:|: :|  :  ::
Db        120 NMPIQGPLEKSLQSSSVSERQRNVEHKVSAIK 151


RESULT 5
A46160
interferon alpha-induced transcription activator ISGF-3, 113K chain - human
N;Alternate names: stat2 protein
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A46160; S71908; S53873
R;Fu, X.Y.; Schindler, C.; Improta, T.; Aebersold, R.; Darnell Jr., J.E.
Proc. Natl. Acad. Sci. U.S.A. 89, 7840-7843, 1992
A;Title: The proteins of ISGF-3, the interferon alpha-induced transcriptional activator,
A;Reference number: A46160; MUID:92366558; PMID:1502204
A;Accession: A46160
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA; protein
A;Residues: 1-851 <FU1>
A;Cross-references: UNIPROT:P52630
A;Note: sequence extracted from NCBI backbone. (NCBIP:110820)
R;Yan, R.; Qureshi, S.; Zhong, Z.; Wen, Z.; Darnell, J.E.
submitted to the EMBL Data Library, December 1994
A;Reference number: S71908
A;Accession: S71908
A;Molecule type: DNA
A;Residues: 1-851 <YAN>
A;Cross-references: EMBL:U18671; NID:g1293919; PIDN:AAA98760.1; PID:g1293920
R;Yan, R.; Qureshi, S.; Zhong, Z.; Wen, Z.; Darnell Jr., J.E.
Nucleic Acids Res. 23, 459-463, 1995
A;Title: The genomic structure of the STAT genes: multiple exons in coincident sites in s
A;Reference number: S53873; MUID:95192056; PMID:7885841
A;Accession: S53873
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-196;392-591;684-730 <YAW>
A;Cross-references: EMBL:U18671
C;Genetics:
A;Gene: stat2
A;Introns: 44/2; 95/3; 127/3; 157/3; 183/1; 211/3; 261/2; 314/2; 345/2; 365/2; 372/2; 40
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: signal transduction; transcription regulation

  Query Match             30.9%;  Score 247.5;  DB 2;  Length 851;
  Best Local Similarity   37.6%;  Pred. No. 7.5e-14;
  Matches   59;  Conservative   34;  Mismatches   55;  Indels    9;  Gaps    6;
```

```
Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDT-FPMELRQFLAPWIESQDWAYAA--SKESHATLVF 57
              |||| || ||: :  |||||||  :  |:::||:|| ||| |:|  ||  | :| ||::|
Db          1 MAQWEMLQNLDSPFQDQLHQLYSHSLLPVDIRQYLAVWIEDQNWQEAALGSDDSKATMLF 60

Qy         58 HNLLGEIDQQYSRFLQE-SNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLL 116
               : | ::: :  |  |:  ::| |||||:  : :|  : : | ::| ::   | || |:|
Db         61 FHFLDQLNYECGRCSQDPESLLLQHNLRKFCRDIQP-FSQDPTQLAEMIFNLLLEEKRIL 119

Qy        117 QTAATAAQQGGQANHPTAAVVTEKQQ-MLEQHLQDVR 152
                 |  | : |:  |      | ||  :|  : |:|
Db        120 IQAQRAQLEQGE---PVLETPVESQQHEIESRILDLR 153


RESULT 6
I49274
mammary gland factor - mouse
N;Alternate names: STAT5 protein homolog p80
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I49274; S54773; S54727
R;Liu, X.; Robinson, G.W.; Gouilleux, F.; Groner, B.; Hennighausen, L.
Proc. Natl. Acad. Sci. U.S.A. 92, 8831-8835, 1995
A;Title: Cloning and expression of Stat5 and an additional homologue (Stat5b) involved i
A;Reference number: I49273; MUID:96004632; PMID:7568026
A;Accession: I49274
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-786 <RES>
A;Cross-references: UNIPROT:P42232; UNIPROT:Q9JKM1; EMBL:U21110; NID:g747973; PIDN:AAC52
R;Mui, A.L.F.; Wakao, H.; O'Farrell, A.M.; Harada, N.; Miyajima, A.
EMBO J. 14, 1166-1175, 1995
A;Title: Interleukin-3, granulocyte-macrophage colony stimulating factor and interleukin
A;Reference number: S54772; MUID:95237198; PMID:7720707
A;Accession: S54773
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-432,'E',434-786 <MUI>
A;Cross-references: EMBL:Z48539; NID:g758635; PIDN:CAA88420.1; PID:g758636
R;Azam, M.; Erdjument-Bromage, H.; Kreider, B.L.; Xia, M.; Quelle, F.; Basu, R.; Saris,
EMBO J. 14, 1402-1411, 1995
A;Title: Interleukin-3 signals through multiple isoforms of Stat5.
A;Reference number: S54725; MUID:95246733; PMID:7537213
A;Accession: S54727
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-432,'E',434-786 <AZA>
C;Genetics:
A;Gene: Stat5b
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             23.5%;  Score 188;  DB 2;  Length 786;
  Best Local Similarity   30.1%;  Pred. No. 1.1e-08;
  Matches   47;  Conservative   26;  Mismatches   75;  Indels    8;  Gaps    2;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYA----ASKESHATLV 56
              || | | |||    | |:  ||   ||:|:| :|: ||||| |        :   || :
Db          1 MAMWIQAQQLQGDALHQMQALYGQHFPIEVRHYLSQWIESQAWDSIDLDNPQENIKATQL 60

Qy         57 FHNLLGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLL 116
                |: |: ::    : |   | :  |      ||| |   |||: | :   |: | ||:
Db         61 LEGLVQELQKKAEHQVGEDGFLLKIKLGHYATQLQSTYDRCPMELVRCIRHILYNEQRLV 120

Qy        117 QTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVR 152
              : |   :   |      | :::|   : |  :::|
Db        121 REANNGSSPAGS----LADAMSQKHLQINQTFEELR 152


RESULT 7
G02317
transcription activator stat5A - human
C;Species: Homo sapiens (man)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 09-Jul-2004
C;Accession: G02317
R;Lin, J.
submitted to the EMBL Data Library, December 1995
A;Reference number: H01043
A;Accession: G02317
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-794 <LIN>
A;Cross-references: UNIPROT:P42229; EMBL:U43185; NID:g1151169; PIDN:AAB06589.1; PID:g115
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             23.4%;  Score 187;  DB 2;  Length 794;
  Best Local Similarity   31.4%;  Pred. No. 1.4e-08;
  Matches   48;  Conservative   23;  Mismatches   68;  Indels   14;  Gaps    3;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYA----ASKESHATLV 56
              || | | |||    |:|:  ||   ||:|:| :|| ||||| |          : || :
Db          1 MAGWIQAQQLQGDALRQMQVLYGQHFPIEVRHYLAQWIESQPWDAIDLDNPQDRAQATQL 60

Qy         57 FHNLLGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLL 116
                |: |: ::    : |   | :  |      ||  |   |:|: | :   |: | ||:
Db         61 LEGLVQELQKKAEHQVGEDGFLLKIKLRHYATQLQKTYDRCPLELVRCIRHILYNEQRLV 120

Qy        117 QTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQ 149
              : |         : |   :|     | ::|||
Db        121 REA-------NNCSSPAGILV---DAMSQKHLQ 143


RESULT 8
S54772
mammary gland factor - mouse
N;Alternate names: stat5 protein
C;Species: Mus musculus (house mouse)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S54772; I49273
R;Mui, A.L.F.; Wakao, H.; O'Farrell, A.M.; Harada, N.; Miyajima, A.
EMBO J. 14, 1166-1175, 1995
A;Title: Interleukin-3, granulocyte-macrophage colony stimulating factor and interleukin
A;Reference number: S54772; MUID:95237198; PMID:7720707
A;Accession: S54772
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-793 <MUI>
A;Cross-references: UNIPROT:P42230; UNIPROT:Q9JIA0; EMBL:Z48538; NID:g758633; PIDN:CAA88
R;Liu, X.; Robinson, G.W.; Gouilleux, F.; Groner, B.; Hennighausen, L.
Proc. Natl. Acad. Sci. U.S.A. 92, 8831-8835, 1995
A;Title: Cloning and expression of Stat5 and an additional homologue (Stat5b) involved i
A;Reference number: I49273; MUID:96004632; PMID:7568026
A;Accession: I49273
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-793 <RES>
A;Cross-references: EMBL:U21103; NID:g747971; PIDN:AAA80590.1; PID:g747972
C;Genetics:
A;Gene: Stat5a
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             23.2%;  Score 186;  DB 2;  Length 793;
  Best Local Similarity   29.5%;  Pred. No. 1.7e-08;
  Matches   46;  Conservative   26;  Mismatches   76;  Indels    8;  Gaps    2;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYA----ASKESHATLV 56
              || | | |||    |:|:  ||   ||:|:| :|| ||||| |          || :
Db          1 MAGWIQAQQLQGDALRQMQVLYGQHFPIEVRHYLAQWIESQPWDAIDLDNPQDRGQATQL 60

Qy         57 FHNLLGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLL 116
                |: |: ::    : |   | :  |      ||: |   |||: | :   |: | ||:
Db         61 LEGLVQELQKKAEHQVGEDGFLLKIKLGHYATQLQNTYDRCPMELVRCIRHILYNEQRLV 120

Qy        117 QTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVR 152
              : |   :   |          :::|   : |  :::|
Db        121 REANNCSSPAG----VLVDAMSQKHLQINQRFEELR 152
```

RESULT 9
S55527
mammary gland factor - sheep
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S55527; S44353
R;Wakao, H.; Gouilleux, F.; Groner, B.
EMBO J. 14, 854-855, 1995
A;Title: Corrigenda. Mammary gland factor (MGF) is a novel member of the cytokine regula
A;Reference number: S55527; MUID:95188889; PMID:7882987
A;Accession: S55527
A;Molecule type: mRNA
A;Residues: 1-794 <WAK>
A;Cross-references: UNIPROT:P42231; EMBL:X78428; NID:g602354; PIDN:CAA55191.1; PID:g6023
A;Note: this is a revision to the sequence from reference S44353
R;Wakao, H.; Gouilleux, F.; Groner, B.
EMBO J. 13, 2182-2191, 1994
A;Title: Mammary gland factor (MGF) is a novel member of the cytokine regulated transcri
A;Reference number: S44353; MUID:94244619; PMID:7514531
A;Accession: S44353
A;Molecule type: mRNA
A;Residues: 1-716,'RHLHGPGSLPSR',729,'P',731,'ASL' <WAW>
A;Cross-references: EMBL:X78428
A;Note: this sequence has been revised in reference S55527
C;Superfamily: human signal transducer and transcription activator STAT5A

Query Match 18.4%; Score 147.5; DB 2; Length 794;
Best Local Similarity 27.3%; Pred. No. 4.1e-05;
Matches 42; Conservative 27; Mismatches 70; Indels 15; Gaps 4;

```
Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYA----ASKESHATLV 56
              ||  |  | |||     |:|:  ||   ||:|:| :|| ||||| |          :  | :
Db          1 MAGWIQAQQLQGDALRQMQVLYGQHFPIEVRHYLAQWIESQPWDAIDLDNPQDRAQVTQL 60

Qy         57 FHNLLGEIDQQYSRFLQESNVLYQHNL-RRIKQFLQSRYLEKPMEIARIVARCLWEESRL 115
                |: |: ::    : |   | :  |   :    ::|    |    | :   |: | ||
Db         61 LEGLVQELQKKAEHQVGEDGFLLKIKLGHYVHVSSRTRTTAAPWSWLRCIRHILYNEQRL 120

Qy        116 LQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQ 149
              :: |          : :| :: :   | ::|||
Db        121 VREATN--------GNSSAGILVD--AMSQKHLQ 144
```

RESULT 10
AI0354
probable lipoprotein YPO2915 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AI0354
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AI0354
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-331 <KUR>
A;Cross-references: UNIPROT:Q8ZCQ6; GB:AL590842; PIDN:CAC92164.1; PID:g15980878; GSPDB:G
C;Genetics:
A;Gene: YPO2915

Query Match 12.2%; Score 98; DB 2; Length 331;
Best Local Similarity 29.3%; Pred. No. 0.32;
Matches 36; Conservative 21; Mismatches 54; Indels 12; Gaps 4;

```
Qy         22 YSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNLLGEIDQQYSRFLQESNVLYQH 81
              || |||: :|  :  | | |     | ::|    : :  |  | | |  | |:|: |  |:
Db        131 YSQTFPVAVRPLIQLWREQQVLRIALAEE---RIRYQRLQDESDAQIDR-LRENQVRLQY 186

Qy         82 NLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQT---AATAAQQGGQANHPTAAVVT 138
              ||        :| ||   :| | ::     :: : :|   | :||:     | | ||
```

Db 187 NL-----LDTTRKLENLTDIERQLSSRKQLQNEIPETDAEAKSAAEAKSAENQPAAAKPA 241

Qy 139 EKQ 141
| :
Db 242 ESK 244

RESULT 11
I57557
DNA-Binding Protein and transcription factor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C;Accession: I57557
R;Quelle, F.W.; Shimoda, K.; Thierfelder, W.; Fischer, C.L.; Kim, A.; Ruben, S.M.; Clevel
Mol. Cell. Biol. 15, 3336-3343, 1995
A;Title: Cloning of murine Stat6 and human Stat6, Stat proteins that are tyrosine phosphc
A;Reference number: I57557; MUID:95280934; PMID:7760829
A;Accession: I57557
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-837 <RES>
A;Cross-references: UNIPROT:P52633; GB:L47650; NID:g1008876; PIDN:AAA79006.1; PID:g10088
C;Genetics:
A;Gene: STAT6
C;Superfamily: human signal transducer and transcription activator STAT5A

Query Match 12.1%; Score 97; DB 2; Length 837;
Best Local Similarity 24.8%; Pred. No. 1.2;
Matches 29; Conservative 21; Mismatches 55; Indels 12; Gaps 4;

```
Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              |: |  : ::     ::| :|| | ||  ||  || |:||| | :    ::    :  |
Db          1 MSLWGLISKMSP---EKLQRLYVD-FPQRLRHLLADWLESQPWEFLVGSDAFCYNMASAL 56

Qy         61 LGEIDQQYSRFLQE---SNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESR 114
              |     |:       |     | :  |         |:| |   |:::   : : |  | :
Db         57 LSATVQRLQATAGEQGKGNSILPH-----ISTLESIYQRDPLKLVATIRQILQGEKK 108
```

RESULT 12
A54740
interleukin-4-induced transcription factor stat - human
C;Species: Homo sapiens (man)
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004
C;Accession: A54740
R;Hou, J.; Schindler, U.; Henzel, W.J.; Ho, T.C.; Brasseur, M.; McKnight, S.L.
Science 265, 1701-1706, 1994
A;Title: An interleukin-4-induced transcription factor: IL-4 stat.
A;Reference number: A54740; MUID:94367369; PMID:8085155
A;Accession: A54740
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-848 <HOU>
A;Cross-references: UNIPROT:P42226
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: DNA binding; transcription regulation

Query Match 11.4%; Score 91.5; DB 2; Length 848;
Best Local Similarity 21.7%; Pred. No. 3.7;
Matches 43; Conservative 25; Mismatches 69; Indels 61; Gaps 8;

```
Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              |: |  : ::     ::: :|| | ||  ||  |  |:||| | :    ::    :  |
Db          1 MSLWGLVSKMPP---EKVQRLYVD-FPQHLRHLLGDWLESQPWEFLVGSDAFCCNLASAL 56

Qy         61 LGEIDQ----------QYSRFLQESNVL---YQHN----LRRIKQFLQS---------RY 94
              | : |           : |  ||  : |   || :    :   :| ||           |:
Db         57 LSDTVQHLQASVGEQGEGSTILQHISTLESIYQRDPLKLVATFRQILQGKKKAVMEQFRH 116

Qy         95 LEKPMEIARIVARCLWEESRL----------------------LQTAATAAQQGGQANHP 132
              | |           |::  |                      ||  | | |    :
Db        117 LPMPFH---------WKQEELKFKTGLRRLQHRVGEIHLLREALQKGAEAGQVSLHSLIE 167
```

```
Qy        133 TAAVVTEKQQMLEQHLQD 150
              | | |   : |   ||:
Db        168 TPANGTGPSEALAMLLQE 185


RESULT 13
T22552
hypothetical protein ZK1151.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C;Accession: T22552; T27703
R;Harris, B.
submitted to the EMBL Data Library, March 1997
A;Reference number: Z19580
A;Accession: T22552
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1355 <WIL>
A;Cross-references: EMBL:Z92788; PIDN:CAB07214.1; GSPDB:GN00019; CESP:ZK1151.1
A;Experimental source: clone F53B8
R;Harris, B.
submitted to the EMBL Data Library, March 1997
A;Reference number: Z20408
A;Accession: T27703
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1355 <WI2>
A;Cross-references: EMBL:Z93398; PIDN:CAB07724.1; GSPDB:GN00019; CESP:ZK1151.1
A;Experimental source: clone ZK1151
C;Genetics:
A;Gene: CESP:ZK1151.1
A;Map position: 1
A;Introns: 94/3; 124/3; 150/3; 209/3; 248/3; 311/3; 424/3; 734/2; 934/2; 1104/3; 1167/3;

  Query Match             11.0%;  Score 88;  DB 2;  Length 1355;
  Best Local Similarity   25.3%;  Pred. No. 13;
  Matches   42;  Conservative   25;  Mismatches   57;  Indels   42;  Gaps   10;

Qy          4 WNQLQQLD----TRYLKQ-LHQL------YSD-----------------TFPMELRQFLA 35
              |: ::|||     | | || ||::      |:|                 : |||  :
Db        677 WSDIKQLDLPMLTNYYKQLLHEMELREKQYNDVHNQGAALLNQGHPAIRVIEVYLRQMQS 736

Qy         36 PWIESQDWAYAASK--ESHA--TLVFHNLLGEIDQQYSRFLQESNVLYQHNLRRIKQFLQ 91
               |    ||  | ||  | |     |   : : |       ::|| :|  ::|  |    |:
Db        737 QW----DWLLALSKCLEEHLRDALNLKSFMEEASDA-EAWIQEQSVRLENNYNRTDFSLE 791

Qy         92 --SRYLEKPMEIARIVARCLWEESRLLQTAATAA---QQGGQANHP 132
                |:| :  ||  |: :       | :  |: :   |:| :  ||
Db        792 EGERFLRELDEIKEILNKYHQVLMALTERCASISPLWQRGERIPHP 837


RESULT 14
T40287
hypothetical protein SPBC354.07c - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T40287
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Duesterhoeft, A.
submitted to the EMBL Data Library, March 1998
A;Reference number: Z21918
A;Accession: T40287
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-371 <WOO>
A;Cross-references: EMBL:AL022071; PIDN:CAA17807.1; GSPDB:GN00067; SPDB:SPBC354.07c
A;Experimental source: strain 972h-; cosmid c354
C;Genetics:
A;Gene: SPDB:SPBC354.07c
A;Map position: 2
A;Introns: 8/2; 75/3

  Query Match             10.3%;  Score 82.5;  DB 2;  Length 371;
  Best Local Similarity   24.0%;  Pred. No. 8.4;
  Matches   44;  Conservative   31;  Mismatches   57;  Indels   51;  Gaps    9;

Qy          5 NQLQQLDTRYLKQLHQLYSDTFP---MELRQFLAPWIESQDWAYAASKESHATLVFHNLL 61
              ||| ||   :|:: :: | | |   :|   |:||:||     | | :: | : ::
Db        147 NQLGQLRV-HLEKYNETYYITLPNISLEGLWFMAPYIELYGSTYIVSNTNYITKIDYSGR 205

Qy         62 GEIDQQYSRFLQES--NVLYQHN--------------------LRRIKQFLQSRYLE-- 96
              |    | | : |   ::: |                     |:    ||    ||
Db        206 G-----YFRGTKNSFKATIFEKNEDPDYIVEGVWTGESKLTIPSLKSTIFFLSIPSLEAT 260

Qy         97 ----KP-MEIARIVARCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDV 151
                  ||  |:    :| :|:|        :||   |  :     :|: |:  :||  :|:
Db        261 PITVKPESEMGDWESRNVWKE-------VSAALASGNYD-----IVSSKKSTIEQSQRDM 308

Qy        152 RKR 154
              |||
Db        309 RKK 311


RESULT 15
F71884
probable lipopolysaccharide biosynthesis protein - Helicobacter pylori (strain J99)
C;Species: Helicobacter pylori
A;Variety: strain J99
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: F71884
R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A;Reference number: A71800; MUID:99120557; PMID:9923682
A;Accession: F71884
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-373 <ARN>
A;Cross-references: UNIPROT:Q9ZKW4; GB:AE001511; GB:AE001439; NID:g4155382; PIDN:AAD0639
A;Experimental source: strain J99
C;Genetics:
A;Gene: jhp0820

  Query Match             10.3%;  Score 82.5;  DB 2;  Length 373;
  Best Local Similarity   33.8%;  Pred. No. 8.4;
  Matches   27;  Conservative   14;  Mismatches   26;  Indels   13;  Gaps    5;

Qy         18 LHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNLLGEIDQQYSRFLQESNV 77
              ||  |:|  |   :  |  |:  |::|   | | |  :||| ||:|:|:     | :
Db        287 LH-FYADEKPW--KHFGYPY--SKEWHQVAFKTSFESLVFENLVGKIE----TFTE---- 333

Qy         78 LYQHNLRRIKQFLQSRYLEK 97
              |  ||  :   :|| :|  :|
Db        334 LNDHNKKSFFEFLNTRLNKK 353


Search completed: May  5, 2005, 15:01:50
Job time : 18.2329 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 5, 2005, 14:55:35 ; Search time 41.6706 Seconds
(without alignments)
1892.467 Million cell updates/sec

Title: US-10-090-185-8
Perfect score: 800
Sequence: 1 MAQWNQLQQLDTRYLKQLHQ.....AVVTEKQQMLEQHLQDVRKR 154

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	792	99.0	770	1	STA3_HUMAN	P40763 homo sapien
2	792	99.0	770	1	STA3_MOUSE	P42227 mus musculu
3	792	99.0	770	1	STA3_RAT	P52631 rattus norv
4	786	98.2	770	1	STA3_BOVIN	P61635 bos taurus
5	779	97.4	771	2	Q6DV79	Q6dv79 gallus gall
6	760	95.0	769	2	Q9PVX8	Q9pvx8 xenopus lae
7	749	93.6	766	2	Q7ZXK3	Q7zxk3 xenopus lae
8	694.5	86.8	414	2	Q7ZTS5	Q7zts5 brachydanio
9	694.5	86.8	786	2	Q6NV46	Q6nv46 brachydanio
10	694.5	86.8	806	2	O93599	O93599 brachydanio
11	691.5	86.4	767	2	O13133	O13133 oncorhynchu
12	685.5	85.7	765	2	Q6DVF3	Q6dvf3 oryzias lat
13	685.5	85.7	785	2	Q6GUE7	Q6gue7 oryzias lat
14	681.5	85.2	764	2	Q90Y16	Q90y16 tetraodon f
15	401	50.1	751	2	Q8JGN0	Q8jgn0 xenopus lae
16	380	47.5	750	1	STA1_HUMAN	P42224 homo sapien
17	380	47.5	750	2	Q68D00	Q68d00 homo sapien
18	379	47.4	712	2	Q6P6Q7	Q6p6q7 rattus norv
19	379	47.4	1165	2	Q7TP57	Q7tp57 rattus norv
20	378	47.2	712	2	Q99K94	Q99k94 mus musculu
21	378	47.2	749	1	STA1_MOUSE	P42225 mus musculu
22	378	47.2	749	2	Q8C3V4	Q8c3v4 mus musculu
23	378	47.2	749	2	Q8C497	Q8c497 mus musculu
24	378	47.2	749	2	Q9D323	Q9d323 mus musculu
25	378	47.2	749	2	Q9QXK0	Q9qxk0 rattus norv
26	378	47.2	755	2	Q8C8M3	Q8c8m3 mus musculu
27	378	47.2	757	2	Q764M5	Q764m5 sus scrofa
28	354.5	44.3	758	2	Q90Y17	Q90y17 tetraodon f
29	351	43.9	754	2	O13132	O13132 oncorhynchu
30	346	43.2	754	2	O13131	O13131 oncorhynchu
31	336.5	42.1	718	2	Q801Y2	Q801y2 carassius a
32	336.5	42.1	749	2	O93598	O93598 brachydanio
33	336.5	42.1	749	2	Q6P943	Q6p943 brachydanio
34	334.5	41.8	749	1	STA4_MOUSE	P42228 mus musculu
35	333.5	41.7	748	1	STA4_HUMAN	Q14765 homo sapien
36	333.5	41.7	748	2	Q66HB2	Q66hb2 rattus norv
37	331	41.4	652	2	Q7ZZ77	Q7zz77 brachydanio
38	331	41.4	657	2	Q8AW24	Q8aw24 brachydanio
39	331	41.4	667	2	Q8AW20	Q8aw20 brachydanio
40	328	41.0	651	2	Q7ZZ53	Q7zz53 brachydanio
41	324.5	40.6	204	2	Q801X2	Q801x2 brachydanio
42	324.5	40.6	229	2	Q803U9	Q803u9 brachydanio
43	324.5	40.6	528	2	Q8JFU8	Q8jfu8 brachydanio
44	316.5	39.6	122	2	Q8AW27	Q8aw27 brachydanio
45	309	38.6	737	2	Q90Y15	Q90y15 tetraodon f

ALIGNMENTS

RESULT 1
STA3_HUMAN
ID STA3_HUMAN STANDARD; PRT; 770 AA.
AC P40763; O14916; Q9BW54;
DT 01-FEB-1995 (Rel. 31, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Signal transducer and activator of transcription 3 (Acute-phase
DE response factor).
GN Name=STAT3; Synonyms=APRF;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Placenta;
RX MEDLINE=94208062; PubMed=7512451; DOI=10.1016/0092-8674(94)90235-6;
RA Akira S., Nishio Y., Inoue M., Wang X.-J., Wei S., Matsusaka T.,
RA Yoshida K., Sudo T., Naruto M., Kishimoto T.;
RT "Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-
RT related transcription factor involved in the gp130-mediated signaling
RT pathway.";
RL Cell 77:63-71(1994).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=98296260; PubMed=9630560; DOI=10.1016/S0378-1119(98)00185-1;
RA Della Pietra L., Bressan A., Pezzotti A., Serlupi-Crescenzi O.;
RT "Highly conserved amino-acid sequence between murine STAT3 and a
RT revised human STAT3 sequence.";
RL Gene 213:119-124(1998).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM DEL-701), AND VARIANT ILE-143.
RA Rieder M.J., Daniels R.L., da Ponte S.H., Hastings N.C., Ahearn M.O.,
RA Rajkumar N., Yi Q., Nickerson D.A.;
RT "SeattleSNPs. NHLBI HL66682 program for genomic applications, UW-
RT FHCRC, Seattle, WA (URL: http://pga.gs.washington.edu).";
RL Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND DEL-701).
RC TISSUE=Kidney, and Pancreas;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,

```
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   SEQUENCE OF 564-704 FROM N.A.
RC   TISSUE=Liver;
RA   Della Pietra L., Bressan A., Pezzotti A.R., Serlupi-Crescenzi O.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   PHOSPHORYLATION ON SERINE.
RX   MEDLINE=95215843; PubMed=7701321;
RA   Zhang X., Blenis J., Li H.-C., Schindler C., Chen-Kiang S.;
RT   "Requirement of serine phosphorylation for formation of STAT-promoter
RT   complexes.";
RL   Science 267:1990-1994(1995).
RN   [7]
RP   INTERACTION WITH NCOA1.
RX   PubMed=11773079; DOI=10.1074/jbc.M111486200;
RA   Giraud S., Bienvenu F., Avril S., Gascan H., Heery D.M., Coqueret O.;
RT   "Functional interaction of STAT3 transcription factor with the
RT   coactivator NcoA/SRC1a.";
RL   J. Biol. Chem. 277:8004-8011(2002).
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes.
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P40763-1; Sequence=Displayed;
CC       Name=Del-701;
CC         IsoId=P40763-2; Sequence=VSP_010474;
CC   -!- TISSUE SPECIFICITY: Heart, brain, placenta, lung, liver, skeletal
CC       muscle, kidney and pancreas.
CC   -!- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,
CC       LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation
CC       is important for the formation of stable DNA-binding STAT3
CC       homodimers and maximal transcriptional activity.
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L29277; AAA58374.1; -.
DR   EMBL; AJ012463; CAA10032.1; -.
DR   EMBL; AY572796; AAS66986.1; -.
DR   EMBL; BC000627; AAH00627.1; -.
DR   EMBL; BC014482; AAH14482.1; -.
DR   EMBL; AF029311; AAB84254.1; -.
DR   PIR; A54444; A54444.
DR   HSSP; P42227; 1BG1.
DR   TRANSFAC; T01493; -.
DR   Genew; HGNC:11364; STAT3.
DR   H-InvDB; HIX0013840; -.
DR   MIM; 102582; -.
DR   GO; GO:0005737; C:cytoplasm; TAS.
DR   GO; GO:0005634; C:nucleus; TAS.
DR   GO; GO:0005062; F:hematopoietin/interferon-class (D200-domain. . .; TAS.
DR   GO; GO:0003700; F:transcription factor activity; TAS.
DR   GO; GO:0006928; P:cell motility; TAS.
DR   GO; GO:0007259; P:JAK-STAT cascade; TAS.
DR   GO; GO:0000122; P:negative regulation of transcription from P. . .; TAS.
DR   GO; GO:0007399; P:neurogenesis; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
KW   Activator; Alternative splicing; DNA-binding; Nuclear protein;
KW   Phosphorylation; Polymorphism; SH2 domain; Transcription regulation.
FT   DOMAIN      580    670       SH2.
FT   MOD_RES     705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES     727    727       Phosphoserine (By similarity).
FT   VARSPLIC    701    701       Missing (in isoform Del-701).
FT                                /FTId=VSP_010474.
FT   VARIANT      32     32       Q -> K (in dbSNP:1803125).
FT                                /FTId=VAR_018683.
FT   VARIANT     143    143       M -> I.
FT                                /FTId=VAR_018679.
FT   CONFLICT    288    288       Q -> H (in Ref. 1).
FT   CONFLICT    460    460       P -> S (in Ref. 1).
FT   CONFLICT    548    548       K -> N (in Ref. 1).
FT   CONFLICT    561    561       F -> Y (in Ref. 1).
FT   CONFLICT    667    667       V -> L (in Ref. 1).
FT   CONFLICT    730    730       T -> A (in Ref. 1).
SQ   SEQUENCE   770 AA;  88067 MW;  6C00632211C8012D CRC64;

  Query Match             99.0%;  Score 792;  DB 1;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 7.3e-60;
  Matches  152;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              |||||||||||||||:|||||||:||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 2
STA3_MOUSE
ID   STA3_MOUSE     STANDARD;      PRT;   770 AA.
AC   P42227;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Signal transducer and activator of transcription 3 (Acute-phase
DE   response factor).
GN   Name=Stat3; Synonyms=Aprf;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A), AND SEQUENCE OF 154-158; 181-185
RP   AND 632-640.
RC   STRAIN=BALB/c; TISSUE=Liver;
RX   MEDLINE=94208062; PubMed=7512451; DOI=10.1016/0092-8674(94)90235-6;
RA   Akira S., Nishio Y., Inoue M., Wang X.J., Shi W., Matsusaka T.,
RA   Yoshida K., Sudo T., Naruto M., Kishimoto T.;
RT   "Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-
RT   related transcription factor involved in the gp130-mediated signaling
```

```
RT   pathway.";
RL   Cell 77:63-71(1994).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   TISSUE=Thymus;
RX   MEDLINE=94188718; PubMed=8140422;
RA   Zhong Z., Wen Z., Darnell J.E. Jr.;
RT   "Stat3: a STAT family member activated by tyrosine phosphorylation in
RT   response to epidermal growth factor and interleukin-6.";
RL   Science 264:95-98(1994).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   TISSUE=Brain;
RX   MEDLINE=95014185; PubMed=7523373;
RA   Raz R., Durbin J.E., Levy D.E.;
RT   "Acute phase response factor and additional members of the interferon-
RT   stimulated gene factor 3 family integrate diverse signals from
RT   cytokines, interferons, and growth factors.";
RL   J. Biol. Chem. 269:24391-24395(1994).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3B).
RC   STRAIN=BALB/c, and C57BL/6; TISSUE=Liver;
RX   MEDLINE=96016116; PubMed=7568080;
RA   Schaefer T.S., Sanders L.K., Nathans D.;
RT   "Cooperative transcriptional activity of Jun and Stat3 beta, a short
RT   form of Stat3.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:9097-9101(1995).
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   STRAIN=129/SvJ;
RX   PubMed=11161808; DOI=10.1006/geno.2000.6433;
RA   Miyoshi K., Cui Y., Riedlinger G., Lehoczky J., Zon L., Oka T.,
RA   Dewar K., Hennighausen L.;
RT   "Structure of the mouse stat 3/5 locus: evolution from Drosophila to
RT   zebrafish to mouse.";
RL   Genomics 71:150-155(2001).
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   STRAIN=C57BL/6J, and NOD/LtJ;
RA   Davoodi-Semiromi A., She J.-X.;
RT   "A mutant Stat5b with weaker DNA binding defines a key defective
RT   pathway in non-obese diabetic (NOD) mice.";
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   STRAIN=FVB/N; TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [8]
RP   PHOSPHORYLATION SITE SER-727, AND MUTAGENESIS.
RX   MEDLINE=95354205; PubMed=7543024; DOI=10.1016/0092-8674(95)90311-9;
RA   Wen Z., Zhong Z., Darnell J.E. Jr.;
RT   "Maximal activation of transcription by Stat1 and Stat3 requires both
RT   tyrosine and serine phosphorylation.";
RL   Cell 82:241-250(1995).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 136-716.
RX   MEDLINE=98334373; PubMed=9671298; DOI=10.1038/28101;
RA   Becker S., Groner B., Mueller C.W.;
RT   "Three-dimensional structure of the Stat3beta homodimer bound to
RT   DNA.";
RL   Nature 394:145-151(1998).
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes. STAT3B interacts with the N-terminal
CC       part of JUN to activate such promoters in a cooperative way.
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=Stat3A;
CC         IsoId=P42227-1; Sequence=Displayed;
CC       Name=Stat3B;
CC         IsoId=P42227-2; Sequence=VSP_006287;
CC       Name=Del-701;
CC         IsoId=P42227-3; Sequence=VSP_010475;
CC   -!- TISSUE SPECIFICITY: STAT3A is seen in the liver, spleen, and
CC       kidney. STAT3B is also detected in the liver, although in a much
CC       less abundant manner.
CC   -!- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,
CC       LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation
CC       is important for the formation of stable DNA-binding STAT3
CC       homodimers and maximal transcriptional activity (By similarity).
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L29278; AAA37254.1; -.
DR   EMBL; U06922; AAA19452.1; -.
DR   EMBL; U08378; AAA56668.1; -.
DR   EMBL; U30709; AAC52612.1; -.
DR   EMBL; AF246978; AAL59017.1; -.
DR   EMBL; AY299489; AAQ75418.1; -.
DR   EMBL; AY299490; AAQ75419.1; -.
DR   EMBL; BC003806; AAH03806.1; -.
DR   PIR; I49508; I49508.
DR   PDB; 1BG1; X-ray; A=1-722.
DR   TRANSFAC; T01574; -.
DR   MGD; MGI:103038; Stat3.
DR   GO; GO:0005737; C:cytoplasm; IDA.
DR   GO; GO:0005634; C:nucleus; IDA.
DR   GO; GO:0005886; C:plasma membrane; IDA.
DR   GO; GO:0003677; F:DNA binding; IDA.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   GO; GO:0016563; F:transcriptional activator activity; IDA.
DR   GO; GO:0007259; P:JAK-STAT cascade; IDA.
DR   GO; GO:0006357; P:regulation of transcription from Pol II pro. . .; IDA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
KW   3D-structure; Activator; Acute phase; Alternative splicing;
KW   Direct protein sequencing; DNA-binding; Nuclear protein;
```

```
KW   Phosphorylation; SH2 domain; Transcription regulation.
FT   DOMAIN       580    670       SH2.
FT   MOD_RES      705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES      727    727       Phosphoserine.
FT   VARSPLIC     716    770       TTCSNTIDLPMSPRTLDSLMQFGNNGEGAEPSAGGQFESLT
FT                                 FDMDLTSECATSPM -> FIDAVWK (in isoform
FT                                 Stat3B).
FT                                 /FTId=VSP_006287.
FT   VARSPLIC     701    701       Missing (in isoform Del-701).
FT                                 /FTId=VSP_010475.
FT   MUTAGEN      727    727       S->A: Decreased transcriptional
FT                                 activation.
FT   CONFLICT      16     16       E -> K (in Ref. 2).
FT   CONFLICT      25     25       S -> T (in Ref. 2 and 4).
FT   CONFLICT     394    394       M -> I (in Ref. 1).
FT   HELIX        139    180
FT   TURN         181    182
FT   TURN         197    198
FT   HELIX        199    237
FT   TURN         238    238
FT   HELIX        239    251
FT   TURN         252    253
FT   HELIX        261    290
FT   TURN         294    295
FT   TURN         297    301
FT   HELIX        302    320
FT   STRAND       321    328
FT   TURN         330    331
FT   TURN         333    334
FT   TURN         336    337
FT   STRAND       338    340
FT   TURN         341    342
FT   STRAND       345    351

  Query Match             99.0%;  Score 792;  DB 1;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 7.3e-60;
  Matches  152;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 3
STA3_RAT
ID   STA3_RAT       STANDARD;      PRT;   770 AA.
AC   P52631;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Signal transducer and activator of transcription 3.
GN   Name=Stat3;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=96102059; PubMed=8530402; DOI=10.1074/jbc.270.50.29998;
RA   Ripperger J.A., Fritz S., Richter K., Hocke G.M., Lottspeich F.,
RA   Fey G.H.;
RT   "Transcription factors Stat3 and Stat5b are present in rat liver
RT   nuclei late in an acute phase response and bind interleukin-6 response
RT   elements.";
RL   J. Biol. Chem. 270:29998-30006(1995).
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes.
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation (By similarity).
CC   -!- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,
CC       LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation
CC       is important for the formation of stable DNA-binding STAT3
CC       homodimers and maximal transcriptional activity (By similarity).
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X91810; CAA62920.1; -.
DR   HSSP; P42227; 1BG1.
DR   RGD; 3772; Stat3.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
KW   Activator; DNA-binding; Nuclear protein; Phosphorylation; SH2 domain;
KW   Transcription regulation.
FT   DOMAIN       580    670       SH2.
FT   MOD_RES      705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES      727    727       Phosphoserine (By similarity).
SQ   SEQUENCE   770 AA;  88039 MW;  D74A0C76954754ED CRC64;

  Query Match             99.0%;  Score 792;  DB 1;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 7.3e-60;
  Matches  152;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||:||||||||:||||||||||||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 4
STA3_BOVIN
ID   STA3_BOVIN     STANDARD;      PRT;   770 AA.
AC   P61635;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Signal transducer and activator of transcription 3.
GN   Name=STAT3;
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
```

```
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Mammary gland;
RA   Seyfert H.M., Wheeler T.T., Moolenaar A., Pitra C.;
RT   "The STAT5B-encoding gene was flipped across the STAT3/STAT5A-locus
RT   during ruminant evolution.";
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes (By similarity).
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation (By similarity).
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ620655; CAF06182.1; -.
DR   PROSITE; PS50001; SH2; 1.
KW   Activator; DNA-binding; Nuclear protein; Phosphorylation; SH2 domain;
KW   Transcription regulation.
FT   DOMAIN      580    670       SH2.
FT   MOD_RES     705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES     727    727       Phosphoserine (By similarity).
SQ   SEQUENCE   770 AA;  87974 MW;  9CEB147C73E83274 CRC64;

  Query Match             98.2%;  Score 786;  DB 1;  Length 770;
  Best Local Similarity   98.1%;  Pred. No. 2.4e-59;
  Matches  151;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              |||||||||||||||:||||||||:||||||| |||||||||||||||||||||||||||
Db          1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQLLAPWIESQDWAYAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              |||||||||||||||||||||||||||||||||
Db        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 5
Q6DV79
ID   Q6DV79      PRELIMINARY;      PRT;   771 AA.
AC   Q6DV79;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Signal transducer and activator of transcription 3.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zhou G.Y., Leung F.C.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY641397; AAT64887.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   771 AA;  88174 MW;  71AC855C5DEC03E2 CRC64;

  Query Match             97.4%;  Score 779;  DB 2;  Length 771;
  Best Local Similarity   96.8%;  Pred. No. 9.7e-59;
  Matches  149;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              |||||||||||||||:|||||| |:||||||||||||||||||:||||||||||||
Db          1 MAQWNQLQQLDTRYLEQLHQLYGDSFPMELRQFLAPWIESQDWAYAANKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              ||||||||| |||||||||||||||||||||||
Db        121 TAAQQGGQATHPTAAVVTEKQQMLEQHLQDVRKR 154


RESULT 6
Q9PVX8
ID   Q9PVX8      PRELIMINARY;      PRT;   769 AA.
AC   Q9PVX8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat 3.
GN   Name=stat 3;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20107399; PubMed=10642787; DOI=10.1006/dbio.1999.9518;
RA   Nishinakamura R., Matsumoto Y., Matsuda T., Ariizumi T., Heike T.,
RA   Asashima M., Yokota T.;
RT   "Activation of Stat3 by cytokine receptor gp130 ventralizes Xenopus
RT   embryos independent of BMP-4.";
RL   Dev. Biol. 216:481-490(1999).
DR   EMBL; AB017701; BAA86061.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   769 AA;  87974 MW;  0905C03263303069 CRC64;

  Query Match             95.0%;  Score 760;  DB 2;  Length 769;
  Best Local Similarity   94.2%;  Pred. No. 4.2e-57;
  Matches  145;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||:||||||||:||||||||||||||||||:||||||||||||||||
Db          1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAFAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEEGRLLQTAA 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              |||||| |:|| ||||||||||||||||||:
Db        121 AAAQQGGPASHPNAAVVTEKQQMLEQHLQDVRKK 154


RESULT 7
Q7ZXK3
ID   Q7ZXK3      PRELIMINARY;      PRT;   766 AA.
AC   Q7ZXK3;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat3-A protein.
GN   Name=stat3-A;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC044717; AAH44717.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   766 AA;  87599 MW;  31018A3321CCEB9C CRC64;

  Query Match             93.6%;  Score 749;  DB 2;  Length 766;
  Best Local Similarity   92.2%;  Pred. No. 3.7e-56;
  Matches  142;  Conservative    6;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||||||||:||||||||:||||||||||||||||||:||||||||||||||||
Db          1 MAQWNQLQQLDTRYLEQLHQLYSDSFPMELRQFLAPWIESQDWAFAASKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              |||||||||||||||||||||||||||||||| |||||||||||||||||| ||||||
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSTYLEKPMEIARIVARCLWEEGRLLQTAV 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              |||||| |:|| |||:||||||||||||||:
Db        121 AAAQQGGPASHPNAAVITEKQQMLEQHLQDVRKK 154


RESULT 8
Q7ZTS5
ID   Q7ZTS5      PRELIMINARY;      PRT;   414 AA.
AC   Q7ZTS5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat3 protein.
GN   Name=stat3;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB; TISSUE=Whole body;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB; TISSUE=Whole body;
RA   Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC045276; AAH45276.1; -.
DR   HSSP; P42227; 1BG1.
```

```
DR   ZFIN; ZDB-GENE-980526-68; stat3.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
SQ   SEQUENCE   414 AA;  48253 MW;  0FFD1B509B7526BD CRC64;

  Query Match             86.8%;  Score 694.5;  DB 2;  Length 414;
  Best Local Similarity   86.4%;  Pred. No. 9.3e-52;
  Matches  133;  Conservative   10;  Mismatches   10;  Indels    1;  Gaps    1;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||:||||:||: ||||:||||||||||||||||||:||||||||||||
Db          1 MAQWNQLQQLETRYLEQLYHLYSDSFPMELRQFLAPWIESQDWAYAANKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              |||||||||||||:|||||||||||||| |||:|||||||||||||||||||| ||||| |
Db         61 LGEIDQQYSRFLQENNVLYQHNLRRIKQHLQSKYLEKPMEIARIVARCLWEEQRLLQT-A 119

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              | ||| ||  |||  |||||||:|| :|||:|||
Db        120 TTAQQDGQVAHPTGTVVTEKQQILEHNLQDIRKR 153


RESULT 9
Q6NV46
ID   Q6NV46      PRELIMINARY;      PRT;   786 AA.
AC   Q6NV46;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Stat3 protein.
GN   Name=stat3;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC068320; AAH68320.1; -.
DR   ZFIN; ZDB-GENE-980526-68; stat3.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   786 AA;  90039 MW;  FC7371D0B0E5447E CRC64;

  Query Match             86.8%;  Score 694.5;  DB 2;  Length 786;
  Best Local Similarity   86.4%;  Pred. No. 2e-51;
  Matches  133;  Conservative   10;  Mismatches   10;  Indels    1;  Gaps    1;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||:||||:||: ||||:||||||||||||||||||:||||||||||||
Db          1 MAQWNQLQQLETRYLEQLYHLYSDSFPMELRQFLAPWIESQDWAYAANKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              |||||||||||||:|||||||||||||| |||:|||||||||||||||||||| ||||| |
Db         61 LGEIDQQYSRFLQENNVLYQHNLRRIKQHLQSKYLEKPMEIARIVARCLWEEQRLLQT-A 119

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              | ||| ||  |||  |||||||:|| :|||:|||
Db        120 TTAQQDGQVAHPTGTVVTEKQQILEHNLQDIRKR 153


RESULT 10
O93599
ID   O93599      PRELIMINARY;      PRT;   806 AA.
AC   O93599;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Transcription factor.
GN   Name=stat3;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Oates A.C.;
RL   Thesis (1998), University of Melbourne, Australia.
DR   EMBL; AJ005693; CAA06677.1; -.
DR   HSSP; P42227; 1BG1.
DR   ZFIN; ZDB-GENE-980526-68; stat3.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   806 AA;  92151 MW;  74BC4EA401C3C942 CRC64;

  Query Match             86.8%;  Score 694.5;  DB 2;  Length 806;
  Best Local Similarity   86.4%;  Pred. No. 2e-51;
```

```
  Matches  133;  Conservative   10;  Mismatches   10;  Indels    1;  Gaps    1;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||:||||:||: ||||:||||||||||||||||||||:||||||||||||
Db          1 MAQWNQLQQLETRYLEQLYHLYSDSFPMELRQFLAPWIESQDWAYAANKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              |||||||||||||:|||||||||||| |||:|||||||||||||||||| ||||| |
Db         61 LGEIDQQYSRFLQENNVLYQHNLRRIKQHLQSKYLEKPMEIARIVARCLWEEQRLLQT-A 119

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              | ||| ||  |||  |||||||:|| :|||:|||
Db        120 TTAQQDGQVAHPTGTVVTEKQQILEHNLQDIRKR 153


RESULT 11
O13133
ID   O13133      PRELIMINARY;.     PRT;   767 AA.
AC   O13133;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat3.
GN   Name=rbtStat3;
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8022;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Johnson M.C., Mourich D.V., Leong J.C.;
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U60333; AAB60926.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
SQ   SEQUENCE   767 AA;  87816 MW;  FED97740C74C3798 CRC64;

  Query Match             86.4%;  Score 691.5;  DB 2;  Length 767;
  Best Local Similarity   85.7%;  Pred. No. 3.4e-51;
  Matches  132;  Conservative   12;  Mismatches    9;  Indels    1;  Gaps    1;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||:||||:||: ||||:||||||||||||||||||||:||||||||||||
Db          1 MAQWNQLQQLETRYLEQLYHLYSDSFPMELRQFLAPWIESQDWAYAANKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              |||||||||||||:|||||||||||| |||:|||||||||||||||||| ||||| |
Db         61 LGEIDQQYSRFLQENNVLYQHNLRRIKQHLQSKYLEKPMEIARIVARCLWEEQRLLQTAT 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              |||| | |:||:  |||||||:|| :|||:|||
Db        121 TAAQD-GTASHPSGTVVTEKQQILEHNLQDIRKR 153


RESULT 12
Q6DVF3
ID   Q6DVF3      PRELIMINARY;      PRT;   765 AA.
AC   Q6DVF3;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Signal transducer and activation of transcription factor 3.
OS   Oryzias latipes (Medaka fish) (Japanese ricefish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Liu R., Hong Y.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY639947; AAT64912.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   765 AA;  87566 MW;  F5D01408748EC703 CRC64;

  Query Match             85.7%;  Score 685.5;  DB 2;  Length 765;
  Best Local Similarity   85.1%;  Pred. No. 1.1e-50;
  Matches  131;  Conservative   11;  Mismatches   11;  Indels    1;  Gaps    1;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||:||||:||: ||||:||||||||||||||||||||:||||||||||||
Db          1 MAQWNQLQQLETRYLEQLYHLYSDSFPMELRQFLAPWIESQDWAYAANKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              |||||||||||||:|||||||||||| |||:|||||||||||||||||| ||||| |
Db         61 LGEIDQQYSRFLQENNVLYQHNLRRIKQHLQSKYLEKPMEIARIVARCLWEEQRLLQT-A 119

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              |:  | | | |||  |||||||:|| :|||:|||
Db        120 TSTSQDGSAAHPTGTVVTEKQQILEHNLQDIRKR 153


RESULT 13
Q6GUE7
ID   Q6GUE7      PRELIMINARY;      PRT;   785 AA.
AC   Q6GUE7;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Signal transducer and activator of transcription 3 isoform 1.
OS   Oryzias latipes (Medaka fish) (Japanese ricefish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Liu R., Hong Y.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY641434; AAT46364.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
```

```
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   785 AA;  89643 MW;  81F231BDE27DE938 CRC64;

  Query Match             85.7%;  Score 685.5;  DB 2;  Length 785;
  Best Local Similarity   85.1%;  Pred. No. 1.2e-50;
  Matches  131;  Conservative   11;  Mismatches   11;  Indels    1;  Gaps    1;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||:||||:||: ||||:||||||||||||||||||||:||||||||||||
Db          1 MAQWNQLQQLETRYLEQLYHLYSDSFPMELRQFLAPWIESQDWAYAANKESHATLVFHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              |||||||||||||:|||||||||||||| |||:|||||||||||||||||||| ||||| |
Db         61 LGEIDQQYSRFLQENNVLYQHNLRRIKQHLQSKYLEKPMEIARIVARCLWEEQRLLQT-A 119

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              |:  | | | |||  |||||||:|| :|||:|||
Db        120 TSTSQDGSAAHPTGTVVTEKQQILEHNLQDIRKR 153


RESULT 14
Q90Y16
ID   Q90Y16      PRELIMINARY;      PRT;   764 AA.
AC   Q90Y16;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   STAT3.
OS   Tetraodon fluviatilis (Puffer fish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=47145;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sung S.C., Fan T.J., Leu J.H., Chou C.M., Huang C.J.;
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF307106; AAL09415.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   764 AA;  87411 MW;  E661FFE18BEFD8BE CRC64;

  Query Match             85.2%;  Score 681.5;  DB 2;  Length 764;
  Best Local Similarity   85.7%;  Pred. No. 2.5e-50;
  Matches  132;  Conservative   11;  Mismatches   10;  Indels    1;  Gaps    1;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||||||||:||||:||: ||||:||||||||||||||||||||:|||||||| |||
Db          1 MAQWNQLQQLETRYLEQLYHLYSDSFPMELRQFLAPWIESQDWAYAANKESHATLVSHNL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              ||||||||||||||||||||||||| |||:|||||||||||||||||| ||||||:
Db         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQHLQSKYLEKPMEIARIVARCLWEEQRLLQTAS 120

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
              | ||  ||| :|:  |||||||||| :|||:|||
Db        121 TVAQD-GQAANPSGTVVTEKQQMLELNLQDIRKR 153


RESULT 15
Q8JGN0
ID   Q8JGN0      PRELIMINARY;      PRT;   751 AA.
AC   Q8JGN0;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat1.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21098508; PubMed=11164887; DOI=10.1016/S0145-305X(00)00050-1;
RA   Turpen J.B., Carlson D.L., Huang C.;
RT   "Cloning and developmental expression of Xenopus Stat1.";
RL   Dev. Comp. Immunol. 25:219-229(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Turpen J.B., Carlson D.L., Huang C.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY101602; AAM51552.1; -.
DR   HSSP; P42224; 1BF5.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   751 AA;  86497 MW;  788810A08B0889EA CRC64;

  Query Match             50.1%;  Score 401;  DB 2;  Length 751;
  Best Local Similarity   48.7%;  Pred. No. 4e-26;
  Matches   75;  Conservative   33;  Mismatches   44;  Indels    2;  Gaps    1;

Qy          1 MAQWNQLQQLDTRYLKQLHQLYSDTFPMELRQFLAPWIESQDWAYAASKESHATLVFHNL 60
              ||||  |||:||::|:|:|||| |:||||:||:|| |:|:||| :||:  | ||::|| |
Db          1 MAQWYDLQQIDTKFLEQVHQLYDDSFPMEIRQYLAQWLENQDWEHAANNVSVATILFHGL 60

Qy         61 LGEIDQQYSRFLQESNVLYQHNLRRIKQFLQSRYLEKPMEIARIVARCLWEESRLLQTAA 120
              | ::| |||||  |:| | |||:|: |: ||  : | |:::| |:  || |||::|  |
Db         61 LSQLDDQYSRFSLENNFLLQHNIRKSKRNLQDHFQEDPVQMAMIINNCLREESKILCNA- 119

Qy        121 TAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKR 154
                  |     |       |: :||: |:: : ||| |
Db        120 -RASNKNQVGSTQTTVMLDKQKELDKKVTDVRNR 152


Search completed: May  5, 2005, 14:57:40
Job time : 43.6706 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 5, 2005, 14:59:00 ; Search time 114.139 Seconds
(without alignments)
792.064 Million cell updates/sec

Title: US-10-090-185-9
Perfect score: 1388
Sequence: 1 RCLWEESRLLQTAATAAQQG.....LNYQLKIKVCIDKDSGDVAA 271

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1428581 seqs, 333598853 residues

Total number of hits satisfying chosen parameters: 1428581

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1388	100.0	271	13	US-10-090-185-9	Sequence 9, Appli
2	1388	100.0	770	11	US-09-876-773-12	Sequence 12, Appl
3	1388	100.0	770	17	US-10-639-617-12	Sequence 12, Appl
4	1377	99.2	720	15	US-10-380-020-4	Sequence 4, Appli
5	1377	99.2	769	15	US-10-380-020-2	Sequence 2, Appli
6	1377	99.2	769	15	US-10-380-020-5	Sequence 5, Appli
7	1377	99.2	770	14	US-10-045-792-8	Sequence 8, Appli
8	1377	99.2	770	14	US-10-038-010-56	Sequence 56, Appl
9	1377	99.2	770	14	US-10-117-087-2	Sequence 2, Appli
10	1377	99.2	793	9	US-09-925-302-780	Sequence 780, App
11	1377	99.2	793	10	US-09-925-302-780	Sequence 780, App
12	1374	99.0	770	15	US-10-116-275-329	Sequence 329, App
13	1372	98.8	770	15	US-10-116-275-349	Sequence 349, App
14	1290	92.9	252	13	US-10-090-185-14	Sequence 14, Appl
15	1212	87.3	236	13	US-10-090-185-15	Sequence 15, Appl
16	1172	84.4	229	13	US-10-090-185-18	Sequence 18, Appl
17	1172	84.4	229	13	US-10-090-185-28	Sequence 28, Appl
18	1168	84.1	229	13	US-10-090-185-31	Sequence 31, Appl
19	1167	84.1	229	13	US-10-090-185-30	Sequence 30, Appl
20	1155	83.2	229	13	US-10-090-185-29	Sequence 29, Appl
21	1142	82.3	223	13	US-10-090-185-22	Sequence 22, Appl
22	1094	78.8	213	13	US-10-090-185-19	Sequence 19, Appl
23	943	67.9	185	13	US-10-090-185-23	Sequence 23, Appl
24	899	64.8	176	13	US-10-090-185-16	Sequence 16, Appl
25	723	52.1	143	13	US-10-090-185-17	Sequence 17, Appl
26	669	48.2	129	13	US-10-090-185-24	Sequence 24, Appl
27	654	47.1	749	9	US-09-833-205-4	Sequence 4, Appli
28	653	47.0	128	13	US-10-090-185-20	Sequence 20, Appl
29	647	46.6	268	13	US-10-090-185-12	Sequence 12, Appl
30	647	46.6	712	11	US-09-876-773-6	Sequence 6, Appli
31	647	46.6	712	14	US-10-245-120-2	Sequence 2, Appli
32	647	46.6	712	17	US-10-639-617-6	Sequence 6, Appli
33	647	46.6	712	17	US-10-936-390-5	Sequence 5, Appli
34	647	46.6	750	9	US-09-833-205-2	Sequence 2, Appli
35	647	46.6	750	11	US-09-876-773-4	Sequence 4, Appli
36	647	46.6	750	14	US-10-245-120-1	Sequence 1, Appli
37	647	46.6	750	14	US-10-308-279-44	Sequence 44, Appl
38	647	46.6	750	16	US-10-755-889-352	Sequence 352, App
39	647	46.6	750	16	US-10-755-889-823	Sequence 823, App
40	647	46.6	750	17	US-10-492-043-19	Sequence 19, Appl
41	647	46.6	750	17	US-10-639-617-4	Sequence 4, Appli
42	647	46.6	786	9	US-09-925-297-550	Sequence 550, App
43	623	44.9	582	14	US-10-245-120-3	Sequence 3, Appli
44	619	44.6	749	9	US-09-833-205-6	Sequence 6, Appli
45	619	44.6	749	11	US-09-876-773-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-10-090-185-9
; Sequence 9, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/10/090,185
; CURRENT FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: 09/387,418
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-090-185-9

Query Match 100.0%; Score 1388; DB 13; Length 271;
Best Local Similarity 100.0%; Pred. No. 2.7e-110;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
```

```
Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271


RESULT 2
US-09-876-773-12
; Sequence 12, Application US/09876773
; Publication No. US20040058318A1
;   GENERAL INFORMATION:
;        APPLICANT: Darnell Jr., James E.
;                   Schindler, Christian W.
;                   Fu, Xian-Yuan
;                   Wen, Zilong
;                   Zhong, Zhong
;        TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
;                            SEQUENCES AND METHODS OF USE THEREOF
;        NUMBER OF SEQUENCES: 25
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Klauber & Jackson
;             STREET: 411 Hackensack Avenue
;             CITY: Hackensack
;             STATE: New Jersey
;             COUNTRY: USA
;             ZIP: 07601
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/876,773
;             FILING DATE: 07-Jun-2001
;             CLASSIFICATION: 435
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/212,185
;             FILING DATE: 11-MAR-1994
;             APPLICATION NUMBER: US 07/980,498
;             FILING DATE: 23-NOV-1992
;             APPLICATION NUMBER: US 07/854,296
;             FILING DATE: 19-MAR-1992
;             APPLICATION NUMBER: WO US93/02569
;             FILING DATE: 19-MAR-1993
;             APPLICATION NUMBER: US 08/126,588
;             FILING DATE: 24-SEP-1993
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Jackson Esq., David A.
;             REGISTRATION NUMBER: 26,742
;             REFERENCE/DOCKET NUMBER: 600-1-073 CIP
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 201 487-5800
;             TELEFAX: 201 343-1684
;             TELEX: 133521
;   INFORMATION FOR SEQ ID NO: 12:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 770 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-876-773-12

  Query Match             100.0%;  Score 1388;  DB 11;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 1e-109;
  Matches  271;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 3
US-10-639-617-12
; Sequence 12, Application US/10639617
; Publication No. US20050079543A1
;   GENERAL INFORMATION:
;        APPLICANT: Darnell Jr., James E.
;                   Schindler, Christian W.
;                   Fu, Xian-Yuan
;                   Wen, Zilong
;                   Zhong, Zhong
;        TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
;                            SEQUENCES AND METHODS OF USE THEREOF
;        NUMBER OF SEQUENCES: 25
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Klauber & Jackson
;             STREET: 411 Hackensack Avenue
;             CITY: Hackensack
;             STATE: New Jersey
;             COUNTRY: USA
;             ZIP: 07601
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/639,617
;             FILING DATE: 12-Aug-2003
;             CLASSIFICATION: 435
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/212,185
;             FILING DATE: 11-MAR-1994
;             APPLICATION NUMBER: US 07/980,498
;             FILING DATE: 23-NOV-1992
;             APPLICATION NUMBER: US 07/854,296
;             FILING DATE: 19-MAR-1992
;             APPLICATION NUMBER: WO US93/02569
;             FILING DATE: 19-MAR-1993
;             APPLICATION NUMBER: US 08/126,588
;             FILING DATE: 24-SEP-1993
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Jackson Esq., David A.
;             REGISTRATION NUMBER: 26,742
;             REFERENCE/DOCKET NUMBER: 600-1-073 CIP
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 201 487-5800
;             TELEFAX: 201 343-1684
;             TELEX: 133521
;   INFORMATION FOR SEQ ID NO: 12:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 770 amino acids
```

```
;             TYPE: amino acid
;             TOPOLOGY: linear
;          MOLECULE TYPE: protein
;          SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-639-617-12

  Query Match             100.0%;  Score 1388;  DB 17;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 1e-109;
  Matches  271;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              ||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 4
US-10-380-020-4
; Sequence 4, Application US/10380020
; Publication No. US20040052762A1
; GENERAL INFORMATION:
;  APPLICANT: Yu, Hua
;  APPLICANT:  Pardoll, Drew
;  APPLICANT:  Jove, Richard
;  APPLICANT:  Dalton, Willian
;  TITLE OF INVENTION: Stat3 Agonists and Antagonists and Therapeutic Uses Thereof
;  FILE REFERENCE: 10873-009-999
;  CURRENT APPLICATION NUMBER: US/10/380,020
;  CURRENT FILING DATE:  2003-03-07
;  PRIOR APPLICATION NUMBER: 60/231,212
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
;   LENGTH: 720
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-380-020-4

  Query Match             99.2%;  Score 1377;  DB 15;  Length 720;
  Best Local Similarity   99.3%;  Pred. No. 8.3e-109;
  Matches  269;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              |||||||||||||||||||| :||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              ||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 5
US-10-380-020-2
; Sequence 2, Application US/10380020
; Publication No. US20040052762A1
; GENERAL INFORMATION:
;  APPLICANT: Yu, Hua
;  APPLICANT:  Pardoll, Drew
;  APPLICANT:  Jove, Richard
;  APPLICANT:  Dalton, Willian
;  TITLE OF INVENTION: Stat3 Agonists and Antagonists and Therapeutic Uses Thereof
;  FILE REFERENCE: 10873-009-999
;  CURRENT APPLICATION NUMBER: US/10/380,020
;  CURRENT FILING DATE:  2003-03-07
;  PRIOR APPLICATION NUMBER: 60/231,212
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 769
;   TYPE: PRT<
;   ORGANISM: Homo Sapiens
US-10-380-020-2

  Query Match             99.2%;  Score 1377;  DB 15;  Length 769;
  Best Local Similarity   99.3%;  Pred. No. 9.1e-109;
  Matches  269;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              |||||||||||||||||||| :||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              ||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 6
US-10-380-020-5
; Sequence 5, Application US/10380020
; Publication No. US20040052762A1
; GENERAL INFORMATION:
;  APPLICANT: Yu, Hua
;  APPLICANT:  Pardoll, Drew
;  APPLICANT:  Jove, Richard
;  APPLICANT:  Dalton, Willian
;  TITLE OF INVENTION: Stat3 Agonists and Antagonists and Therapeutic Uses Thereof
;  FILE REFERENCE: 10873-009-999
;  CURRENT APPLICATION NUMBER: US/10/380,020
;  CURRENT FILING DATE:  2003-03-07
;  PRIOR APPLICATION NUMBER: 60/231,212
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 769
```

```
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-380-020-5

  Query Match             99.2%;  Score 1377;  DB 15;  Length 769;
  Best Local Similarity   99.3%;  Pred. No. 9.1e-109;
  Matches  269;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              |||||||||||||||||| :||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 7
US-10-045-792-8
; Sequence 8, Application US/10045792
; Publication No. US20030003563A1
;   GENERAL INFORMATION:
;        APPLICANT: Vinkemeier, Uwe
;                   Moarefi, Ismail
;                   Darnell, Jr., James E.
;                   Kuriyan, John
;        TITLE OF INVENTION: A CRYSTAL OF THE N-TERMINAL DOMAIN OF A
;                            STAT PROTEIN AND METHODS OF USE THEREOF
;        NUMBER OF SEQUENCES: 13
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Klauber & Jackson
;             STREET: 411 Hackensack Avenue, 4th Floor
;             CITY: Hackensack
;             STATE: New Jersey
;             COUNTRY: USA
;             ZIP: 07601
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/045,792
;             FILING DATE: 19-Oct-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/012,710
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Jackson Esq., David A.
;             REGISTRATION NUMBER: 26,742
;             REFERENCE/DOCKET NUMBER: 600-1-194
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 201-487-5800
;             TELEFAX: 201-343-1684
;             TELEX: 133521
;   INFORMATION FOR SEQ ID NO: 8:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 770 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        HYPOTHETICAL: NO
;        SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-045-792-8

  Query Match             99.2%;  Score 1377;  DB 14;  Length 770;
  Best Local Similarity   99.3%;  Pred. No. 9.1e-109;
  Matches  269;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              |||||||||||||||||| :||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 8
US-10-038-010-56
; Sequence 56, Application US/10038010
; Publication No. US20030040089A1
; GENERAL INFORMATION:
;  APPLICANT: HYBRIGENICS
;  APPLICANT:  Pierre, Legrain
;  TITLE OF INVENTION: Protein-protein interactions in adipocyte cells
;  FILE REFERENCE: B4767A
;  CURRENT APPLICATION NUMBER: US/10/038,010
;  CURRENT FILING DATE:  2002-07-23
;  PRIOR APPLICATION NUMBER: US 60/259,377
;  PRIOR FILING DATE: 2001-01-02
;  NUMBER OF SEQ ID NOS: 67
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: STAT3 : Transcription factor
;   LOCATION: (1)..(770)
;   OTHER INFORMATION:
US-10-038-010-56

  Query Match             99.2%;  Score 1377;  DB 14;  Length 770;
  Best Local Similarity   99.3%;  Pred. No. 9.1e-109;
  Matches  269;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              |||||||||||||||||| :||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
                 |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 9
US-10-117-087-2
; Sequence 2, Application US/10117087
; Publication No. US20030166854A1
; GENERAL INFORMATION:
;  APPLICANT: SERLUPI-CRESCENZI, Ottaviano
;  APPLICANT:  DELLA PIETRA, Linda
;  TITLE OF INVENTION: ALLELIC VARIANTS OF HUMAN STAT3
;  FILE REFERENCE: SERLUPI=2
;  CURRENT APPLICATION NUMBER: US/10/117,087
;  CURRENT FILING DATE:  2002-04-08
;  PRIOR APPLICATION NUMBER: US/09/526,542
;  PRIOR FILING DATE: 2000-03-19
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Human
US-10-117-087-2

  Query Match             99.2%;  Score 1377;  DB 14;  Length 770;
  Best Local Similarity   99.3%;  Pred. No. 9.1e-109;
  Matches  269;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
                 |||||||||||||||||| :|||||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
                 |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 10
US-09-925-302-780
; Sequence 780, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA104
;  CURRENT APPLICATION NUMBER: US/09/925,302
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05918
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 896
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 780
;   LENGTH: 793
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-925-302-780

  Query Match             99.2%;  Score 1377;  DB 9;  Length 793;
  Best Local Similarity   99.3%;  Pred. No. 9.4e-109;
  Matches  269;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 189

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 249

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
                 |||||||||||||||||| :|||||||||||||||||||||||||||||||||||||||||||
Db        250 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 309

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 369

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
                 |||||||||||||||||||||||||||||||
Db        370 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 400


RESULT 11
US-09-925-302-780
; Sequence 780, Application US/09925302
; Publication No. US20030064072A9
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA104
;  CURRENT APPLICATION NUMBER: US/09/925,302
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05918
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 896
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 780
;   LENGTH: 793
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-925-302-780

  Query Match             99.2%;  Score 1377;  DB 10;  Length 793;
  Best Local Similarity   99.3%;  Pred. No. 9.4e-109;
  Matches  269;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 189

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 249

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
                 |||||||||||||||||| :|||||||||||||||||||||||||||||||||||||||||||
Db.       250 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 309

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 369

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
                 |||||||||||||||||||||||||||||||
Db        370 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 400
```

RESULT 12
US-10-116-275-329
; Sequence 329, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elan Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; TITLE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell Receptors
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 329
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Rat
US-10-116-275-329

Query Match 99.0%; Score 1374; DB 15; Length 770;
Best Local Similarity 98.9%; Pred. No. 1.6e-108;
Matches 268; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              |||||||||||||||||||| :|||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377
```

RESULT 13
US-10-116-275-349
; Sequence 349, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elan Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; TITLE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell Receptors
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 349
; LENGTH: 770
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-275-349

Query Match 98.8%; Score 1372; DB 15; Length 770;
Best Local Similarity 98.9%; Pred. No. 2.4e-108;
Matches 268; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              |||||||||||||||||||| :|||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LHQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377
```

RESULT 14
US-10-090-185-14
; Sequence 14, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR PI
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/10/090,185
; CURRENT FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: 09/387,418
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-090-185-14

Query Match 92.9%; Score 1290; DB 13; Length 252;
Best Local Similarity 100.0%; Pred. No. 5.7e-102;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240

Qy        241 TKVRLLVKFPEL 252
              ||||||||||||
```

```
Db        241 TKVRLLVKFPEL 252


RESULT 15
US-10-090-185-15
; Sequence 15, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
;   LENGTH: 236
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-15

  Query Match             87.3%;  Score 1212;  DB 13;  Length 236;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-95;
  Matches  236;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTG 236
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTG 236


Search completed: May  5, 2005, 15:17:41
Job time : 115.139 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 5, 2005, 14:55:35 ; Search time 40.8094 Seconds
(without alignments)
495.716 Million cell updates/sec

Title: US-10-090-185-9
Perfect score: 1388
Sequence: 1 RCLWEESRLLQTAATAAQQG.....LNYQLKIKVCIDKDSGDVAA 271

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1388	100.0	271	3	US-09-387-418A-9	Sequence 9, Appli
2	1388	100.0	770	1	US-08-369-796-12	Sequence 12, Appl
3	1388	100.0	770	2	US-08-852-091-12	Sequence 12, Appl
4	1388	100.0	770	2	US-08-820-754-12	Sequence 12, Appl
5	1388	100.0	770	3	US-08-956-652-12	Sequence 12, Appl
6	1388	100.0	770	3	US-08-956-869-12	Sequence 12, Appl
7	1388	100.0	770	3	US-08-948-547-12	Sequence 12, Appl
8	1388	100.0	770	3	US-09-364-970-3	Sequence 3, Appli
9	1388	100.0	770	3	US-09-364-970-5	Sequence 5, Appli
10	1388	100.0	770	3	US-08-956-653A-12	Sequence 12, Appl
11	1388	100.0	770	4	US-08-212-185-12	Sequence 12, Appl
12	1388	100.0	770	5	PCT-US95-17025-12	Sequence 12, Appl
13	1377	99.2	770	1	US-08-416-581B-9	Sequence 9, Appli
14	1377	99.2	770	3	US-09-012-710-8	Sequence 8, Appli
15	1377	99.2	770	3	US-09-556-273-8	Sequence 8, Appli
16	1377	99.2	770	3	US-09-526-542-2	Sequence 2, Appli
17	1377	99.2	770	4	US-10-117-087-2	Sequence 2, Appli
18	1372	98.8	770	1	US-08-416-581B-1	Sequence 1, Appli
19	1372	98.8	770	1	US-08-416-581B-5	Sequence 5, Appli
20	1372	98.8	770	3	US-09-087-465-6	Sequence 6, Appli
21	1372	98.8	770	4	US-09-972-800A-6	Sequence 6, Appli
22	1372	98.8	771	1	US-08-276-099A-14	Sequence 14, Appl
23	1372	98.8	771	1	US-08-781-890-14	Sequence 14, Appl
24	1290	92.9	252	3	US-09-387-418A-14	Sequence 14, Appl
25	1212	87.3	236	3	US-09-387-418A-15	Sequence 15, Appl
26	1172	84.4	229	3	US-09-387-418A-18	Sequence 18, Appl
27	1172	84.4	229	3	US-09-387-418A-28	Sequence 28, Appl
28	1168	84.1	229	3	US-09-387-418A-31	Sequence 31, Appl
29	1167	84.1	229	3	US-09-387-418A-30	Sequence 30, Appl
30	1155	83.2	229	3	US-09-387-418A-29	Sequence 29, Appl
31	1142	82.3	223	3	US-09-387-418A-22	Sequence 22, Appl
32	1094	78.8	213	3	US-09-387-418A-19	Sequence 19, Appl
33	943	67.9	185	3	US-09-387-418A-23	Sequence 23, Appl
34	899	64.8	176	3	US-09-387-418A-16	Sequence 16, Appl
35	723	52.1	143	3	US-09-387-418A-17	Sequence 17, Appl
36	669	48.2	129	3	US-09-387-418A-24	Sequence 24, Appl
37	653	47.0	128	3	US-09-387-418A-20	Sequence 20, Appl
38	647	46.6	268	3	US-09-387-418A-12	Sequence 12, Appl
39	647	46.6	712	1	US-08-369-796-6	Sequence 6, Appli
40	647	46.6	712	2	US-08-852-091-6	Sequence 6, Appli
41	647	46.6	712	2	US-08-820-754-6	Sequence 6, Appli
42	647	46.6	712	3	US-08-956-652-6	Sequence 6, Appli
43	647	46.6	712	3	US-08-956-869-6	Sequence 6, Appli
44	647	46.6	712	3	US-08-948-547-6	Sequence 6, Appli
45	647	46.6	712	3	US-08-956-653A-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-09-387-418A-9
; Sequence 9, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/09/387,418A
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-387-418A-9

Query Match 100.0%; Score 1388; DB 3; Length 271;
Best Local Similarity 100.0%; Pred. No. 4.2e-119;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
```

RESULT 2
US-08-369-796-12
; Sequence 12, Application US/08369796

```
; Patent No. 5716622
;  GENERAL INFORMATION:
;    APPLICANT:  James E. Darnell, Jr.
;    APPLICANT:  Zilong Wen
;    APPLICANT:  Curt M. Horvath
;    APPLICANT:  Zhong Zhong
;    TITLE OF INVENTION:  FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
;    TITLE OF INVENTION:  TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/369,796
;      FILING DATE:  06-JAN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-116
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-369-796-12

  Query Match           100.0%;  Score 1388;  DB 1;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-118;
  Matches  271;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 3
US-08-852-091-12
; Sequence 12, Application US/08852091
; Patent No. 5883228
;  GENERAL INFORMATION:
;    APPLICANT:  James E. Darnell, Jr.
;    APPLICANT:  Zilong Wen
;    APPLICANT:  Curt M. Horvath
;    APPLICANT:  Zhong Zhong
;    TITLE OF INVENTION:  FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
;    TITLE OF INVENTION:  TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/852,091
;      FILING DATE:  06-MAY-1997
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/369,796
;      FILING DATE:  06-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-116
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-852-091-12

  Query Match           100.0%;  Score 1388;  DB 2;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-118;
  Matches  271;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 4
US-08-820-754-12
; Sequence 12, Application US/08820754
; Patent No. 5976835
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
```

```
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/820,754
;      FILING DATE:  19-MAR-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,185
;      FILING DATE:  11-MAR-1994
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-820-754-12

  Query Match             100.0%;  Score 1388;  DB 2;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-118;
  Matches  271;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 5
US-08-956-652-12
; Sequence 12, Application US/08956652
; Patent No. 6013475
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/956,652
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,185
;      FILING DATE:  11-MAR-1994
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-956-652-12

  Query Match             100.0%;  Score 1388;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-118;
  Matches  271;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226
```

```
Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377
```

RESULT 6
US-08-956-869-12
; Sequence 12, Application US/08956869
; Patent No. 6030808
; GENERAL INFORMATION:
; APPLICANT: Darnell Jr., James E.
; APPLICANT: Schindler, Christian W.
; APPLICANT: Fu, Xián-Yuan
; APPLICANT: Wen, Zilong
; APPLICANT: Zhong, Zhong
; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
; TITLE OF INVENTION: SEQUENCES AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,869
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/212,185
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/854,296
; FILING DATE: 19-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO US93/02569
; FILING DATE: 19-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,588
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-073 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-956-869-12

Query Match 100.0%; Score 1388; DB 3; Length 770;
Best Local Similarity 100.0%; Pred. No. 1.7e-118;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377
```

RESULT 7
US-08-948-547-12
; Sequence 12, Application US/08948547
; Patent No. 6124118
; GENERAL INFORMATION:
; APPLICANT: Darnell Jr., James E.
; APPLICANT: Schindler, Christian W.
; APPLICANT: Fu, Xian-Yuan
; APPLICANT: Wen, Zilong
; APPLICANT: Zhong, Zhong
; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
; TITLE OF INVENTION: SEQUENCES AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,547
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/212,185
; FILING DATE: 11-MAR-1994
; APPLICATION NUMBER: US 07/980,498
; FILING DATE: 23-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/854,296
; FILING DATE: 19-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO US93/02569
; FILING DATE: 19-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,588
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-073 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684

```
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-948-547-12

  Query Match           100.0%;  Score 1388;  DB 3;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-118;
  Matches  271;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 8
US-09-364-970-3
; Sequence 3, Application US/09364970
; Patent No. 6235873
; GENERAL INFORMATION:
;  APPLICANT: Bromberg, Jacqueline
;  TITLE OF INVENTION: CONSTITUTIVELY ACTIVE TRANSCRIPTION FACTORS AND THEIR
;  TITLE OF INVENTION: USES FOR IDENTIFYING MODULATORS OF ACTIVITY INCLUDING
;  TITLE OF INVENTION: DYSPROLIFERATIVE CELLULAR CHANGES
;  FILE REFERENCE: 600-1-252
;  CURRENT APPLICATION NUMBER: US/09/364,970
;  CURRENT FILING DATE: 1999-07-31
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-364-970-3

  Query Match           100.0%;  Score 1388;  DB 3;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-118;
  Matches  271;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 9
US-09-364-970-5
; Sequence 5, Application US/09364970
; Patent No. 6235873
; GENERAL INFORMATION:
;  APPLICANT: Bromberg, Jacqueline
;  TITLE OF INVENTION: CONSTITUTIVELY ACTIVE TRANSCRIPTION FACTORS AND THEIR
;  TITLE OF INVENTION: USES FOR IDENTIFYING MODULATORS OF ACTIVITY INCLUDING
;  TITLE OF INVENTION: DYSPROLIFERATIVE CELLULAR CHANGES
;  FILE REFERENCE: 600-1-252
;  CURRENT APPLICATION NUMBER: US/09/364,970
;  CURRENT FILING DATE: 1999-07-31
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-364-970-5

  Query Match           100.0%;  Score 1388;  DB 3;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-118;
  Matches  271;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 10
US-08-956-653A-12
; Sequence 12, Application US/08956653A
; Patent No. 6338949
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
```

```
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/956,653A
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/212,185
;       FILING DATE:  11-MAR-1994
;       APPLICATION NUMBER:  US 07/980,498
;       FILING DATE:  23-NOV-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/854,296
;       FILING DATE:  19-MAR-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  WO US93/02569
;       FILING DATE:  19-MAR-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/126,588
;       FILING DATE:  24-SEP-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Jackson Esq., David A.
;       REGISTRATION NUMBER:  26,742
;       REFERENCE/DOCKET NUMBER:  600-1-195
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  201 487-5800
;       TELEFAX:  201 343-1684
;       TELEX:  133521
;   INFORMATION FOR SEQ ID NO:  12:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  770 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-956-653A-12

  Query Match             100.0%;  Score 1388;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-118;
  Matches  271;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 11
US-08-212-185-12
; Sequence 12, Application US/08212185
; Patent No. 6605442
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/212,185
;      FILING DATE:  11-MAR-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;   INFORMATION FOR SEQ ID NO:  12:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  770 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-212-185-12

  Query Match             100.0%;  Score 1388;  DB 4;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-118;
  Matches  271;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 12
PCT-US95-17025-12
```

```
; Sequence 12, Application PC/TUS9517025
;  GENERAL INFORMATION:
;    APPLICANT:  James E. Darnell, Jr.
;    APPLICANT:  Zilong Wen
;    APPLICANT:  Curt M. Horvath
;    APPLICANT:  Zhong Zhong
;    TITLE OF INVENTION:  FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
;    TITLE OF INVENTION:  TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/17025
;      FILING DATE:  28-DEC-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/369,796
;      FILING DATE:  06-JAN-1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-116
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US95-17025-12

  Query Match             100.0%;  Score 1388;  DB 5;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-118;
  Matches  271;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 13
US-08-416-581B-9
; Sequence 9, Application US/08416581B
; Patent No. 5719042
;  GENERAL INFORMATION:
;    APPLICANT:  Kishimoto, Tadamitsu
;    APPLICANT:  Akira, Shizuo
;    TITLE OF INVENTION:  TRANSCRIPTION FACTOR APRF
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SUGHRUE, MION, ZINN, MACPEAK & SEAS
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/416,581B
;      FILING DATE:  04-APR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-65825/1994
;      FILING DATE:  04-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Nakamura, Dean H.
;      REGISTRATION NUMBER:  33,981
;      REFERENCE/DOCKET NUMBER:  Q-37891
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)293-7060
;      TELEFAX:  (202)293-7860
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-416-581B-9

  Query Match             99.2%;  Score 1377;  DB 1;  Length 770;
  Best Local Similarity   99.3%;  Pred. No. 1.7e-117;
  Matches  269;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              |||||||||||||||||||| :||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 14
US-09-012-710-8
; Sequence 8, Application US/09012710
; Patent No. 6087478
;  GENERAL INFORMATION:
;    APPLICANT:  Vinkemeier, Uwe
;    APPLICANT:  Moarefi, Ismail
;    APPLICANT:  Darnell, Jr., James E.
```

```
;    APPLICANT:  Kuriyan, John
;    TITLE OF INVENTION:  A CRYSTAL OF THE N-TERMINAL DOMAIN OF A
;    TITLE OF INVENTION:  STAT PROTEIN AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/012,710
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-194
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-09-012-710-8

  Query Match             99.2%;  Score 1377;  DB 3;  Length 770;
  Best Local Similarity   99.3%;  Pred. No. 1.7e-117;
  Matches  269;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              |||||||||||||||||| :||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 15
US-09-556-273-8
; Sequence 8, Application US/09556273
; Patent No. 6312887
;  GENERAL INFORMATION:
;    APPLICANT:  Vinkemeier, Uwe
;    APPLICANT:  Moarefi, Ismail
;    APPLICANT:  Darnell, Jr., James E.
;    APPLICANT:  Kuriyan, John
;    TITLE OF INVENTION:  A CRYSTAL OF THE N-TERMINAL DOMAIN OF A
;    TITLE OF INVENTION:  STAT PROTEIN AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/556,273
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/012,710
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-194
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-09-556-273-8

  Query Match             99.2%;  Score 1377;  DB 3;  Length 770;
  Best Local Similarity   99.3%;  Pred. No. 1.7e-117;
  Matches  269;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              |||||||||||||||||| :||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


Search completed: May  5, 2005, 14:58:52
Job time : 41.8094 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 5, 2005, 14:55:35 ; Search time 84.1694 Seconds
(without alignments)
1245.251 Million cell updates/sec

Title: US-10-090-185-9
Perfect score: 1388
Sequence: 1 RCLWEESRLLQTAATAAQQG.....LNYQLKIKVCIDKDSGDVAA 271

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1388	100.0	271	4	AAY72841	Aay72841 Mouse Sta
2	1388	100.0	770	2	AAR72082	Aar72082 Mouse Sta
3	1388	100.0	770	2	AAW03176	Aaw03176 Mouse STA
4	1377	99.2	720	5	AAE22055	Aae22055 Human Sta
5	1377	99.2	769	5	ABB57164	Abb57164 Mouse isc
6	1377	99.2	769	5	AAE22054	Aae22054 Human Sta
7	1377	99.2	769	5	AAE22056	Aae22056 Human pro
8	1377	99.2	770	2	AAR82995	Aar82995 Mouse liv
9	1377	99.2	770	2	AAY03768	Aay03768 Human STA
10	1377	99.2	770	3	AAB12377	Aab12377 N-termina
11	1377	99.2	770	5	AAE14652	Aae14652 Murine ST
12	1377	99.2	770	5	ABG69497	Abg69497 Human bai
13	1377	99.2	770	6	ABU10476	Abu10476 Mouse STA
14	1377	99.2	770	8	ADN04365	Adn04365 Antipsori
15	1377	99.2	770	8	ADP54789	Adp54789 Human PRO
16	1377	99.2	793	3	AAB58442	Aab58442 Lung canc
17	1374	99.0	770	7	ADD44738	Add44738 Rat Prote
18	1372	98.8	770	2	AAR82993	Aar82993 Human pla
19	1372	98.8	770	4	AAB19964	Aab19964 Human sig
20	1372	98.8	770	5	AAE15174	Aae15174 Human Sta
21	1372	98.8	770	7	ADD44740	Add44740 Human Pro
22	1290	92.9	252	4	AAY72846	Aay72846 Mouse Sta
23	1212	87.3	236	4	AAY72847	Aay72847 Mouse Sta
24	1172	84.4	229	4	AAY72850	Aay72850 Mouse Sta
25	1168	84.1	229	4	AAY72863	Aay72863 Mouse Sta
26	1167	84.1	229	4	AAY72862	Aay72862 Mouse Sta
27	1157.5	83.4	228	4	AAY72861	Aay72861 Mouse Sta
28	1146	82.6	229	4	AAY72860	Aay72860 Mouse Sta
29	1142	82.3	223	4	AAY72854	Aay72854 Mouse Sta
30	1094	78.8	213	4	AAY72851	Aay72851 Mouse Sta
31	943	67.9	185	4	AAY72855	Aay72855 Mouse Sta
32	899	64.8	176	4	AAY72848	Aay72848 Mouse Sta
33	723	52.1	143	4	AAY72849	Aay72849 Mouse Sta
34	669	48.2	129	4	AAY72856	Aay72856 Mouse Sta
35	654	47.1	749	5	AAG78526	Aag78526 Rat STAT-
36	653	47.0	128	4	AAY72852	Aay72852 Mouse Sta
37	647	46.6	268	4	AAY72844	Aay72844 Mouse Sta
38	647	46.6	680	6	ABR59713	Abr59713 Human sig
39	647	46.6	712	2	AAR72079	Aar72079 Human Sta
40	647	46.6	712	2	AAW03170	Aaw03170 Human STA
41	647	46.6	712	2	AAW62995	Aaw62995 Human Sta
42	647	46.6	712	6	ABU04747	Abu04747 Human exp
43	647	46.6	712	6	ABU04735	Abu04735 Human exp
44	647	46.6	712	6	ABU04745	Abu04745 Human exp
45	647	46.6	712	6	ABU04743	Abu04743 Human exp

ALIGNMENTS

RESULT 1
AAY72841
ID AAY72841 standard; protein; 271 AA.
XX
AC AAY72841;
XX
DT 31-MAY-2001 (first entry)
XX
DE Mouse Stat3 protein fragment #2 (107-377 amino acids).
XX
KW Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW cellular transformation; dysproliferative disease; cancer; psoriasis;
KW therapy.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT Region 24. .48
FT /note= "Stat3-c-Jun interaction region 1; corresponds to
FT 130-154 position of Stat3 protein"
FT Region 236. .252
FT /note= "Stat3-c-Jun interaction region 2; corresponds to
FT 342-358 position of Stat3 protein"
XX
PN WO200116605-A2.
XX
PD 08-MAR-2001.
XX
PF 30-AUG-2000; 2000WO-US023822.
XX
PR 31-AUG-1999; 99US-00387418.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI Zhang X, Horvath C, Wrzeszcynska MH, Darnell JE;
XX
DR WPI; 2001-226705/23.
XX
PT Identifying an agent for use in modulating the interaction between
PT transcription factor c-Jun and a Stat3 protein.
XX
PS Claim 65; Page 67-68; 86pp; English.
XX
CC The present sequence is mouse Stat3 protein fragment containing 107-377
CC amino acids of Stat3 protein. This Stat3 fragment showed strong binding
CC to c-Jun protein in the cell extract. The invention relates to methods
CC for identifying interacting regions of transcription factors and methods
CC for identifying agents which modulates the interaction between a

```
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 271 AA;

  Query Match           100.0%;  Score 1388;  DB 4;  Length 271;
  Best Local Similarity  100.0%;  Pred. No. 2.7e-116;
  Matches  271;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271


RESULT 2
AAR72082
ID   AAR72082 standard; protein; 770 AA.
XX
AC   AAR72082;
XX
DT   25-MAR-2003  (revised)
DT   27-SEP-1995  (first entry)
XX
DE   Mouse Stat3 (19sf6).
XX
KW   Signal transducer and activator of transcription; STAT; 19sf6; Stat3;
KW   receptor recognition factor; transcription factor; cellular debilitation;
KW   derangement; dysfunction; interferon-gamma.
XX
OS   Mus sp.
XX
PN   WO9508629-A1.
XX
PD   30-MAR-1995.
XX
PF   26-SEP-1994;   94WO-US010849.
XX
PR   24-SEP-1993;   93US-00126588.
PR   24-SEP-1993;   93US-00126595.
PR   11-MAR-1994;   94US-00212184.
PR   11-MAR-1994;   94US-00212185.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Darnell JE,  Schindler CW,  Shuai K,  Wen Z,  Zhong Z;
XX
DR   WPI; 1995-139598/18.
DR   N-PSDB; AAQ89340.
XX
PT   Receptor recognition factor implicated in transcriptional stimulation of
PT   genes - useful in drug screening assays and/or for treating cellular
PT   debilitations, derangements and/or dysfunctions, etc.
XX
PS   Claim 1; Page 107-110; 160pp; English.
XX
CC   A fragment encoding the human Stat91 protein was used to screen a murine
CC   thymus and spleen cDNA for homologous proteins. A highly homologous gene
CC   (given in AAQ89338) was isolated that encoded a 91 kDa protein (AAR72080)
CC   (Stat1) that was responsive to interferon- gamma. Using a fragment of the
CC   mouse gene as probe, 2 additional members of the 113-91 family of
CC   receptor recognition factor proteins were isolated. The 2 genes (AAQ89339
CC   -40) were cloned in plasmids 13sf1 and 19sf6 and encoded proteins termed
CC   Stat4 (AAR72081) and Stat3 (AAR72082), respectively. (Updated on 25-MAR-
CC   2003 to correct PN field.)
XX
SQ   Sequence 770 AA;

  Query Match           100.0%;  Score 1388;  DB 2;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 1.1e-115;
  Matches  271;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 3
AAW03176
ID   AAW03176 standard; protein; 770 AA.
XX
AC   AAW03176;
XX
DT   24-OCT-1996  (first entry)
XX
DE   Mouse STAT4.
XX
KW   STAT; STAT4; signal transducer and activator of transcription;
KW   DNA binding protein; ligand; receptor; oncogenesis; inflammation;
KW   autoimmune disease; antagonist; therapy.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Domain          398. .508
FT                   /label= DNA_binding_domain
FT                   /note= "Claim 3, page 110"
XX
PN   WO9620954-A2.
XX
PD   11-JUL-1996.
XX
PF   28-DEC-1995;   95WO-US017025.
XX
PR   06-JAN-1995;   95US-00369796.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Darnell JE,  Wen Z,  Horvath CM,  Zhong Z;
XX
DR   WPI; 1996-333941/33.
```

```
DR   N-PSDB; AAT31280.
XX
PT   New STAT protein DNA-binding domain peptide(s) - useful for diagnosing,
PT   preventing or treating cellular dysfunction, e.g. oncogenesis,
PT   inflammation, parasitic disease or autoimmunity.
XX
PS   Disclosure; Page 87-90; 138pp; English.
XX
CC   Mouse signal transducer and activator of transcription (STAT) protein
CC   STAT4 (AAW03176) serves a dual purpose, i.e. signal transduction from
CC   ligand-activated receptor kinase complexes followed by nuclear
CC   translocation and DNA binding to activate transcription. Recombinant
CC   STAT4 can be obtd. using cDNA clone 19sf6 (AAT31278) obtd. from
CC   splenic/thymic cells. STAT4 includes a DNA-binding domain (see also
CC   AAW03167) capable of both receptor recognition and message delivery via
CC   DNA binding in a receptor-ligand specific manner. STAT proteins and their
CC   DNA binding domains (see also AAW03165-75) are useful for screening
CC   antagonists used to inhibit STAT-mediated signal transduction and
CC   activation of transcription
XX
SQ   Sequence 770 AA;

  Query Match             100.0%;  Score 1388;  DB 2;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-115;
  Matches  271;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 4
AAE22055
ID   AAE22055 standard; protein; 720 AA.
XX
AC   AAE22055;
XX
DT   25-JUL-2002  (first entry)
XX
DE   Human Stat3beta protein.
XX
KW   Human; signal transducer and activator of transcription 3; ischaemia;
KW   immune response; Stat3; coronary atherosclerosis; vascular occlusion;
KW   hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;
KW   inflammation; chronic obstructive pulmonary disease; cardiac arrest;
KW   insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;
KW   shock; chronic active hepatitis; adult respiratory distress syndrome;
KW   nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;
KW   Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;
KW   polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;
KW   proliferative disease; Grave's disease; ulcerative colitis; sarcoma;
KW   carcinoma; degenerative disorder; gene therapy; growth deficiency;
KW   cirrhosis; hypoproliferative disorder; lesion; Statbeta.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 713. .714
FT                   /note= "Encoded by ACA CCA TTC"
XX
PN   WO200220032-A1.
XX
PD   14-MAR-2002.
XX
PF   10-SEP-2001; 2001WO-US028254.
XX
PR   08-SEP-2000; 2000US-0231212P.
XX
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (UYSF-) UNIV SOUTH FLORIDA.
XX
PI   Yu H,  Pardoll D,  Jove R,  Dalton W;
XX
DR   WPI; 2002-362218/39.
DR   N-PSDB; AAD35066.
XX
PT   Modulating angiogenesis and an immune response in an individual, for
PT   treating a hypoxic or ischemic condition, comprises administering a
PT   compound that modulates the activity of a signal transducer and activator
PT   of transcription 3.
XX
PS   Disclosure; Page 87-89; 94pp; English.
XX
CC   The invention relates to a method of modulating angiogenesis and immune
CC   response. Method involves administering to an individual a compound that
CC   modulate the activity of signal transducer and activator of transcription
CC   3 (Stat3). Modulating angiogenesis is useful for treating or preventing
CC   hypoxic or ischaemic condition or disorder which is the result of stroke,
CC   ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,
CC   tissue ischaemia in the lower extremities, infarction, trauma, vascular
CC   occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,
CC   chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,
CC   epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest,
CC   nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy
CC   with neovascularisation. Suppressing an immune response is useful for
CC   ameliorating a symptom of an autoimmune disease such as systemic lupus
CC   erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,
CC   Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,
CC   mixed connective tissue disease, primary biliary cirrhosis, pernicious
CC   anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,
CC   gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,
CC   idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's
CC   disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune
CC   infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and
CC   dense deposit disease. The method is useful in preventing or treating
CC   specific proliferative and oncogenic disease which includes sarcomas and
CC   carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,
CC   fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,
CC   hypoproliferative disorders, physical trauma, lesions and wounds. The
CC   method is also used in gene therapy. The present sequence is human
CC   Stat3beta protein
XX
SQ   Sequence 720 AA;

  Query Match             99.2%;  Score 1377;  DB 5;  Length 720;
  Best Local Similarity   99.3%;  Pred. No. 9.6e-115;
  Matches  269;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              |||||||||||||||||| :|||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286
```

```
Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377
```

RESULT 5
ABB57164
ID ABB57164 standard; protein; 769 AA.
XX
AC ABB57164;
XX
DT 07-MAR-2002 (first entry)
XX
DE Mouse ischaemic condition related protein sequence SEQ ID NO:398.
XX
KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease.
XX
OS Mus musculus.
XX
PN WO200188188-A2.
XX
PD 22-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-JP004192.
XX
PR 18-MAY-2000; 2000JP-00145977.
XX
PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
DR WPI; 2002-034733/04.
DR N-PSDB; ABI99454.
XX
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or by
PT determining the expression profile of a gene group comprising these
PT genes.
XX
PS Claim 2; Page 1084-1087; 2690pp; English.
XX
CC The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (ABI99202 to ABI99912, encoding the
CC protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The expression
CC levels or expression profiles produced by these genes are used as an
CC indicator when screening for ischaemic condition-improving drugs or
CC therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR
CC primers for a mouse ischaemic condition related sequence, which are used
CC in the exemplification of the present invention
XX
SQ Sequence 769 AA;

Query Match 99.2%; Score 1377; DB 5; Length 769;
Best Local Similarity 99.3%; Pred. No. 1e-114;
Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              |||||||||||||||||| :|||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377
```

RESULT 6
AAE22054
ID AAE22054 standard; protein; 769 AA.
XX
AC AAE22054;
XX
DT 25-JUL-2002 (first entry)
XX
DE Human Stat3 protein.
XX
KW Human; signal transducer and activator of transcription 3; ischaemia;
KW immune response; Stat3; coronary atherosclerosis; vascular occlusion;
KW hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;
KW inflammation; chronic obstructive pulmonary disease; cardiac arrest;
KW insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;
KW shock; chronic active hepatitis; adult respiratory distress syndrome;
KW nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;
KW Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;
KW polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;
KW proliferative disease; Grave's disease; ulcerative colitis; sarcoma;
KW carcinoma; degenerative disorder; gene therapy; growth deficiency;
KW cirrhosis; hypoproliferative disorder; lesion.
XX
OS Homo sapiens.
XX
PN WO200220032-A1.
XX
PD 14-MAR-2002.
XX
PF 10-SEP-2001; 2001WO-US028254.
XX
PR 08-SEP-2000; 2000US-0231212P.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
PA (UYSF-) UNIV SOUTH FLORIDA.
XX
PI Yu H, Pardoll D, Jove R, Dalton W;
XX
DR WPI; 2002-362218/39.
DR N-PSDB; AAD35065.
XX
PT Modulating angiogenesis and an immune response in an individual, for
PT treating a hypoxic or ischemic condition, comprises administering a
PT compound that modulates the activity of a signal transducer and activator
PT of transcription 3.
XX
PS Disclosure; Page 83-85; 94pp; English.
XX
CC The invention relates to a method of modulating angiogenesis and immune
CC response. Method involves administering to an individual a compound that
CC modulate the activity of signal transducer and activator of transcription
CC 3 (Stat3). Modulating angiogenesis is useful for treating or preventing
CC hypoxic or ischaemic condition or disorder which is the result of stroke,
CC ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,
CC tissue ischaemia in the lower extremities, infarction, trauma, vascular
CC occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,
CC chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,
CC epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest,

CC nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy
CC with neovascularisation. Suppressing an immune response is useful for
CC ameliorating a symptom of an autoimmune disease such as systemic lupus
CC erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,
CC Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,
CC mixed connective tissue disease, primary biliary cirrhosis, pernicious
CC anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,
CC gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,
CC idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's
CC disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune
CC infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and
CC dense deposit disease. The method is useful in preventing or treating
CC specific proliferative and oncogenic disease which includes sarcomas and
CC carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,
CC fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,
CC hypoproliferative disorders, physical trauma, lesions and wounds. The
CC method is also used in gene therapy. The present sequence is human Stat3
CC protein
XX
SQ Sequence 769 AA;

Query Match 99.2%; Score 1377; DB 5; Length 769;
Best Local Similarity 99.3%; Pred. No. 1e-114;
Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              |||||||||||||||||||| :||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377
```

RESULT 7
AAE22056
ID AAE22056 standard; protein; 769 AA.
XX
AC AAE22056;
XX
DT 25-JUL-2002 (first entry)
XX
DE Human protein related to angiogenesis regulation.
XX
KW Human; signal transducer and activator of transcription 3; ischaemia;
KW immune response; Stat3; coronary atherosclerosis; vascular occlusion;
KW hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;
KW inflammation; chronic obstructive pulmonary disease; cardiac arrest;
KW insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;
KW shock; chronic active hepatitis; adult respiratory distress syndrome;
KW nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;
KW Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;
KW polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;
KW proliferative disease; Grave's disease; ulcerative colitis; sarcoma;
KW carcinoma; degenerative disorder; gene therapy; growth deficiency;
KW cirrhosis; hypoproliferative disorder; lesion.
XX
OS Homo sapiens.
XX
PN WO200220032-A1.
XX
PD 14-MAR-2002.
XX
PF 10-SEP-2001; 2001WO-US028254.
XX
PR 08-SEP-2000; 2000US-0231212P.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
PA (UYSF-) UNIV SOUTH FLORIDA.
XX
PI Yu H, Pardoll D, Jove R, Dalton W;
XX
DR WPI; 2002-362218/39.
XX
PT Modulating angiogenesis and an immune response in an individual, for
PT treating a hypoxic or ischemic condition, comprises administering a
PT compound that modulates the activity of a signal transducer and activator
PT of transcription 3.
XX
PS Disclosure; Page 83-85; 94pp; English.
XX
CC The invention relates to a method of modulating angiogenesis and immune
CC response. Method involves administering to an individual a compound that
CC modulate the activity of signal transducer and activator of transcription
CC 3 (Stat3). Modulating angiogenesis is useful for treating or preventing
CC hypoxic or ischaemic condition or disorder which is the result of stroke,
CC ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,
CC tissue ischaemia in the lower extremities, infarction, trauma, vascular
CC occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,
CC chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,
CC epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest,
CC nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy
CC with neovascularisation. Suppressing an immune response is useful for
CC ameliorating a symptom of an autoimmune disease such as systemic lupus
CC erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,
CC Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,
CC mixed connective tissue disease, primary biliary cirrhosis, pernicious
CC anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,
CC gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,
CC idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's
CC disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune
CC infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and
CC dense deposit disease. The method is useful in preventing or treating
CC specific proliferative and oncogenic disease which includes sarcomas and
CC carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,
CC fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,
CC hypoproliferative disorders, physical trauma, lesions and wounds. The
CC method is also used in gene therapy. The present sequence is human
CC protein related to angiogenesis regulation
XX
SQ Sequence 769 AA;

Query Match 99.2%; Score 1377; DB 5; Length 769;
Best Local Similarity 99.3%; Pred. No. 1e-114;
Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              |||||||||||||||||||| :||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
```

```
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377
```

RESULT 8
AAR82995
ID AAR82995 standard; protein; 770 AA.
XX
AC AAR82995;
XX
DT 25-MAR-1996 (first entry)
XX
DE Mouse liver acute phase response factor.
XX
KW Mouse; acute phase response factor; transcription factor; interleukin-6;
KW signal transmission; liver; antibody; antisense; ribozyme;
KW antiinflammatory; antitumor; hypotensive; therapy.
XX
OS Mus musculus.
XX
PN EP676469-A2.
XX
PD 11-OCT-1995.
XX
PF 29-MAR-1995; 95EP-00104670.
XX
PR 04-APR-1994; 94JP-00065825.
XX
PA (KISH/) KISHIMOTO T.
XX
PI Akira S, Kishimoto T;
XX
DR WPI; 1995-346089/45.
DR N-PSDB; AAT05619.
XX
PT New acute phase response factor - for developing inhibitory agents for
PT treating diseases induced by cytokine(s) such as IL-6, e.g. inflammatory
PT diseases.
XX
PS Claim 10; Page 20-22; 31pp; English.
XX
CC The sequence represents a mouse acute phase response factor (APRF), a
CC transcription factor related to signal transmission of interleukin-6 (IL-
CC 6). The protein is encoded by a cDNA, isolated from a mouse liver cDNA
CC library using a polymerase chain reaction product (amplified using
CC primers derived from an IL-6-treated mouse liver peptide) as a probe.
CC APRF-inhibitors, e.g. antibodies, antisense oligonucleotides or
CC ribozymes, may be used to treat diseases induced by IL-6, e.g.
CC inflammatory disease, leukemia, cancer, osteoclasia, pulmonary
CC hypertension, etc
XX
SQ Sequence 770 AA;

```
  Query Match             99.2%;  Score 1377;  DB 2;  Length 770;
  Best Local Similarity   99.3%;  Pred. No. 1e-114;
  Matches  269;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              |||||||||||||||||||| :||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
```

```
              ||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377
```

RESULT 9
AAY03768
ID AAY03768 standard; protein; 770 AA.
XX
AC AAY03768;
XX
DT 11-JUN-1999 (first entry)
XX
DE Human STAT3 allelic variant.
XX
KW Signal Transducer and Activator of Transcription 3; STAT3; allele; IL-6;
KW intracellular transcription factor; interleukin-6; medicament; variant;
KW pharmaceutical; autoimmune disease; inflammatory; human.
XX
OS Homo sapiens.
XX
PN EP905234-A2.
XX
PD 31-MAR-1999.
XX
PF 18-FEB-1998; 98EP-00102774.
XX
PR 16-SEP-1997; 97EP-00116061.
XX
PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
XX
PI Serlupi-Crescenzi O, Della Pietra L;
XX
DR WPI; 1999-192664/17.
DR N-PSDB; AAX29281.
XX
PT New human Signal Transducer and Activator of Transcription 3 (STAT3)
PT allelic variant useful for treatment of autoimmune and inflammatory
PT disease.
XX
PS Claim 2; Page 9-13; 32pp; English.
XX
CC The present sequence represents a predominant allelic variant of human
CC Signal Transducer and Activator of Transcription 3 (STAT3) protein, an
CC intracellular transcription factor which mediates IL-6 signals. The
CC encoding sequence differs from the original published human STAT3 gene
CC sequence (Akira et al., 1994). Host cells comprising a recombinant STAT3
CC DNA molecule can be used for the recombinant expression of the variant.
CC STAT3 protein is useful as a medicament or pharmaceutical composition for
CC treatment of autoimmune or inflammatory diseases
XX
SQ Sequence 770 AA;

```
  Query Match             99.2%;  Score 1377;  DB 2;  Length 770;
  Best Local Similarity   99.3%;  Pred. No. 1e-114;
  Matches  269;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              |||||||||||||||||||| :||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
```

```
                  ||||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377

RESULT 10
AAB12377
ID   AAB12377 standard; peptide; 770 AA.
XX
AC   AAB12377;
XX
DT   08-NOV-2000  (first entry)
XX
DE   N-terminal domain of murine STAT-3 protein.
XX
KW   STAT; signal transducer and activator of transcription; crystal;
KW   drug design; murine.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Region          4. .9
FT                   /label= Alpha helix 1
FT   Region          12. .21
FT                   /label= Alpha helix 2
FT   Region          19. .21
FT                   /label= 3(10) helix of alpha helix 2
FT   Region          28. .33
FT                   /label= Alpha helix 3
FT   Region          35. .40
FT                   /label= Alpha helix 4
FT   Region          43. .47
FT                   /label= Alpha helix 5
FT   Region          50. .73
FT                   /label= Alpha helix 6
FT   Region          77. .96
FT                   /label= Alpha helix 7
FT   Region          99. .119
FT                   /label= Alpha helix 8
XX
PN   US6087478-A.
XX
PD   11-JUL-2000.
XX
PF   23-JAN-1998;   98US-00012710.
XX
PR   23-JAN-1998;   98US-00012710.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Vinkemeier U,  Moarefi I,  Darnell JE,  Kuriyan J;
XX
DR   WPI; 2000-505108/45.
XX
PT   New crystals of an N-terminal fragment of a signal transducer and
PT   activator of transcription that effectively diffracts x-rays, useful for
PT   drug screening and development.
XX
PS   Disclosure; Fig 1; 42pp; English.
XX
CC   The present invention relates to a crystal of an N-terminal fragment of a
CC   signal transducer and activator of transcription (STAT) protein. The
CC   crystal effectively diffracts X-rays, allowing the determination of the
CC   atomic coordinates of the N-terminal domain to a resolution of greater
CC   than 5.0 Angstroms. The present sequence is the N-terminal domain of the
CC   murine STAT 3 protein. The N-terminal domain enables STAT dimers to
CC   interact and bind DNA cooperatively, a mechanism important for gene
CC   activation. The crystals are useful in drug screening and development by
CC   selecting a potential drug by performing rational drug design with the 3-
CC   dimensional structure determined for the crystal
XX
SQ   Sequence 770 AA;

  Query Match             99.2%;  Score 1377;  DB 3;  Length 770;
  Best Local Similarity   99.3%;  Pred. No. 1e-114;
  Matches  269;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              |||||||||||||||||||| :|||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              ||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377

RESULT 11
AAE14652
ID   AAE14652 standard; protein; 770 AA.
XX
AC   AAE14652;
XX
DT   16-JUL-2002  (first entry)
XX
DE   Murine STAT3 protein.
XX
KW   Signal transducer and activator of transcription; STAT3;
KW   drug development; drug discovery; crystal; inflammation; allergy; asthma;
KW   leukaemia; anaemia; neutropaenia; thrombocytopaenia; cancer; obesity;
KW   viral disease; growth retardation; murine.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Domain          1. .130
FT                   /note= "Conserved N-terminal domain of the STAT family"
XX
PN   US6312887-B1.
XX
PD   06-NOV-2001.
XX
PF   24-APR-2000; 2000US-00556273.
XX
PR   23-JAN-1998;   98US-00012710.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Vinkemeier U,  Moarefi I,  Darnell JE,  Kuriyan J;
XX
DR   WPI; 2002-033337/04.
XX
PT   Identifying compounds that bind to signal transducer and activator of
PT   transcription proteins, useful for the production of new drugs.
XX
PS   Example; Col 47-50; 44pp; English.
XX
CC   The invention relates to methods for detecting compounds that bind to
CC   signal transducer and activator of transcription (STAT) proteins for the
CC   discovery and development of new drug compounds based on the structural
CC   properties of the protein crystal. The methods include: identifying a
CC   compound that binds to the N-terminal domain of a STAT protein,
CC   identifying a compound that enhances or diminishes the binding of the
CC   dimeric STAT proteins to each other and/or their nucleic acid binding
CC   site; or identifying a compound that enhances or diminishes the ability
CC   of STAT protein dimers to induce the expression of a gene operably under
```

```
CC   the control of a promoter containing at least two adjacent weak binding
CC   sites for STAT protein dimers. The methods are used for identifying new
CC   drugs. An antagonist of STAT N-terminal dimeric interactions that
CC   inhibits the binding of the STAT dimers to adjacent weak binding sites on
CC   a promoter of a gene, could be useful as drugs in the treatment of
CC   diseases, e.g. inflammation, allergy, asthma and leukaemias. On the other
CC   hand, an agonist of N-terminal dimeric interactions between STAT dimers,
CC   can be used as drugs in the treatment of diseases e.g. anaemia,
CC   neutropaenia, thrombocytopaenia, cancer, obesity, viral diseases and
CC   growth retardation. The present sequence is murine STAT3 protein
XX
SQ   Sequence 770 AA;

  Query Match             99.2%;  Score 1377;  DB 5;  Length 770;
  Best Local Similarity   99.3%;  Pred. No. 1e-114;
  Matches  269;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              |||||||||||||||||| :||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 12
ABG69497
ID   ABG69497 standard; protein; 770 AA.
XX
AC   ABG69497;
XX
DT   21-OCT-2002  (first entry)
XX
DE   Human bait protein STAT3.
XX
KW   Human; yeast two-hybrid assay; adipocyte; bait protein; NIDDM;
KW   non-insulin diabetes mellitus; obesity; selected interacting domain; SID;
KW   protein-protein interaction map; PIM; anorectic; metabolic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200253726-A2.
XX
PD   11-JUL-2002.
XX
PF   28-DEC-2001; 2001WO-EP015423.
XX
PR   02-JAN-2001; 2001US-0259377P.
XX
PA   (HYBR-) HYBRIGENICS.
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Legrain P,  Marullo S,  Jockers R;
XX
DR   WPI; 2002-583612/62.
DR   N-PSDB; ABS51033.
XX
PT   Novel complex of protein-protein interactions in adipocyte cells for
PT   identifying compounds that modulate the protein-protein interactions and
PT   useful for treating obesity and metabolic disorders.
XX
PS   Claim 1; Page 54; 125pp; English.
XX
CC   The invention relates to a complex of protein-protein interactions
CC   (forming a protein-protein interaction map, PIM) in adipocyte cells as
CC   defined in the specification, or polynucleotides in adipocytes encoding
CC   for the polypeptides. Also included are a recombinant cell expressing the
CC   interacting polypeptides and a method of selecting a modulating compound
CC   in adipocyte cells, by cultivating a recombinant host cell on a selective
CC   medium containing a modulating compound and a reporter gene the
CC   expression of which is toxic for the recombinant host cell which is
CC   transformed with two vectors, where the first vector comprises a
CC   polynucleotide encoding a first hybrid polypeptide and DNA binding domain
CC   and the second vector comprising a polynucleotide encoding a second
CC   hybrid polypeptide and an activating domain that activates the toxic
CC   reporter gene, when the first and second hybrid polypeptides interact and
CC   selecting the modulating compound which inhibits the growth of the
CC   recombinant host cell (i.e. using the yeast two-hybrid system). The
CC   complexes are useful for identifying compounds that modulate the protein-
CC   protein interactions and useful for treating obesity and metabolic
CC   disorders e.g. non-insulin dependent diabetes mellitus, NIDDM. The
CC   compound isolated by the method is useful for treating and preventing
CC   obesity or metabolic diseases. The interactions between the proteins of
CC   the complex further define a set of selected interacting domains, SID.
CC   The present sequence represents a member of the protein complex of the
CC   invention, used as the bait protein in the yeast two- hybrid assay
XX
SQ   Sequence 770 AA;

  Query Match             99.2%;  Score 1377;  DB 5;  Length 770;
  Best Local Similarity   99.3%;  Pred. No. 1e-114;
  Matches  269;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              |||||||||||||||||| :||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 13
ABU10476
ID   ABU10476 standard; protein; 770 AA.
XX
AC   ABU10476;
XX
DT   06-AUG-2003  (first entry)
XX
DE   Mouse STAT3 protein.
XX
KW   Mouse; signal transducer and activator of transcription; drug design;
KW   drug screening; STAT-STAT dimer interaction; STAT3.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Region          4. .9
FT                   /label= alpha_helix_1
FT   Region          12. .21
```

```
FT                   /label= alpha_helix_2
FT                   /note= "Residues 19-21 form a 3 helix"
FT   Region          28. .33
FT                   /label= alpha_helix_3
FT   Region          35. .40
FT                   /label= alpha_helix_4
FT   Region          43. .47
FT                   /label= alpha_helix_5
FT   Region          50. .73
FT                   /label= alpha_helix_6
FT                   /note= "Residues 57, 61, 64, 68 and 71 contribute to
FT                   packing of the coiled-coil"
FT   Region          77. .96
FT                   /label= alpha_helix_7
FT                   /note= "Residues 79, 83, 86, 90 and 94 contribute to
FT                   packing of the coiled-coil"
FT   Region          99. .119
FT                   /label= alpha_helix_8
XX
PN   US2003003563-A1.
XX
PD   02-JAN-2003.
XX
PF   19-OCT-2001; 2001US-00045792.
XX
PR   23-JAN-1998;   98US-00012710.
PR   24-APR-2000; 2000US-00556273.
XX
PA   (VINK/) VINKEMEIER U.
PA   (MOAR/) MOAREFI I.
PA   (DARN/) DARNELL J E.
PA   (KURI/) KURIYAN J.
XX
PI   Vinkemeier U,  Moarefi I,  Darnell JE,  Kuriyan J;
XX
DR   WPI; 2003-447354/42.
XX
PT   New crystal having an N-terminal domain of a STAT protein performing X-
PT   ray crystallographic studies, useful for screening drugs that enhance or
PT   inhibit STAT-STAT dimer interactions.
XX
PS   Disclosure; Page 25-26; 46pp; English.
XX
CC   The invention relates to a crystal of an N-terminal domain of signal
CC   transducer and activator of transcription (STAT) protein, where the
CC   crystal effectively diffracts X-rays for the determination of the atomic
CC   coordinates of the N-terminal domain of the STAT protein to a resolution
CC   of greater than 5.0 Angstrom. The methods and compositions are useful for
CC   the design and screening of drugs that enhance or inhibit STAT-STAT dimer
CC   interactions. The present sequence represents the amino acid sequence of
CC   mouse STAT3 protein
XX
SQ   Sequence 770 AA;
```

```
  Query Match             99.2%;  Score 1377;  DB 6;  Length 770;
  Best Local Similarity   99.3%;  Pred. No. 1e-114;
  Matches  269;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              |||||||||||||||||| :||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377
```

```
RESULT 14
ADN04365
ID   ADN04365 standard; protein; 770 AA.
XX
AC   ADN04365;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Antipsoriatic protein sequence #377.
XX
KW   antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO2004028479-A2.
XX
PD   08-APR-2004.
XX
PF   25-SEP-2003; 2003WO-US030907.
XX
PR   25-SEP-2002; 2002US-0414006P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Bodary S,  Clark H,  Jackman J,  Schoenfeld J,  Williams PM,  Wood WI;
PI   Wu TD;
XX
DR   WPI; 2004-305105/28.
DR   N-PSDB; ADN04364.
XX
PT   New PRO nucleic acid or polypeptide, useful for preparing a
PT   pharmaceutical composition for diagnosing or treating psoriasis in a
PT   mammal.
XX
PS   Claim 9; SEQ ID NO 759; 3069pp; English.
XX
CC   The invention relates to novel polynucleotide and polypeptides for
CC   treating psoriasis or a sequence having at least 80% identity to the
CC   above sequences. The nucleic acid is useful for preparing a composition
CC   for diagnosing or treating psoriasis in a mammal. This sequence
CC   corresponds to one of the polypeptides of the invention.
XX
SQ   Sequence 770 AA;
```

```
  Query Match             99.2%;  Score 1377;  DB 8;  Length 770;
  Best Local Similarity   99.3%;  Pred. No. 1e-114;
  Matches  269;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              |||||||||||||||||| :||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377
```

```
RESULT 15
ADP54789
ID   ADP54789 standard; protein; 770 AA.
XX
AC   ADP54789;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Human PRO protein sequence SEQ ID NO:765.
XX
KW   human; PRO; immune related disease; inflammatory immune response;
KW   immune response stimulation; antiallergic; antianaemic; antiarthritic;
KW   antiasthmatic; antidiabetic; antiinflammatory; antipsoriatic;
KW   antirheumatic; antithyroid; CNS; dermatological; gastrointestinal;
KW   haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;
KW   nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;
KW   virucide; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO2004039956-A2.
XX
PD   13-MAY-2004.
XX
PF   28-OCT-2003; 2003WO-US034381.
XX
PR   29-OCT-2002; 2002US-0422472P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Aggarwal S,  Clark H,  Gurney AL,  Schoenfeld J,  Williams PM;
PI   Wood WI,  Wu TD;
XX
DR   WPI; 2004-376182/35.
DR   N-PSDB; ADP54788.
XX
PT   New PRO polynucleotides and polypeptides, useful in useful in diagnosing
PT   and treating an immune related disease, e.g. systemic lupus
PT   erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in
PT   stimulating an immune response.
XX
PS   Claim 1; SEQ ID NO 765; 3009pp; English.
XX
CC   The present invention describes an isolated PRO nucleic acid (I). Also
CC   described: (1) a vector comprising (I); (2) a host cell comprising the
CC   vector of (1); (3) a process for producing a PRO polypeptides; (4) an
CC   isolated PRO polypeptide; (5) a chimeric molecule comprising the
CC   polypeptide of (4) fused to a heterologous amino acid sequence; (6) an
CC   antibody which specifically binds to a polypeptide of (4); (7) a
CC   composition of matter comprising a polypeptide of (4), an agonist or
CC   antagonist of the polypeptide or an antibody that binds to the
CC   polypeptide in combination with a carrier; (8) an article of manufacture
CC   comprising a container, a label on the container and a composition of
CC   matter of (7); (9) a method of treating an immune related disease in a
CC   mammal; (10) a method for determining the presence of a PRO polypeptide
CC   in a sample suspected of having the polypeptide; (11) a method of
CC   diagnosing an immune related disease or an inflammatory immune response
CC   in mammal; (12) a method of identifying a compound that inhibits or
CC   mimics the activity of or expression of a gene encoding a PRO polypeptide
CC   ; and (13) a method of stimulating the immune response in a mammal. The
CC   PRO sequences have antiallergic, antianaemic, antiarthritic,
CC   antiasthmatic, antidiabetic, antiinflammatory, antipsoriatic,
CC   antirheumatic, antithyroid, CNS, dermatological, gastrointestinal,
CC   haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,
CC   nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and
CC   virucide activities, and can be used in gene therapy. The nucleic acid
CC   (I) and the encoded polypeptides, compositions, kits and methods are
CC   useful in diagnosing and treating an immune related disease and in
CC   stimulating an immune response. The present sequence represents a human
CC   PRO protein from the present invention.
XX
SQ   Sequence 770 AA;

  Query Match             99.2%;  Score 1377;  DB 8;  Length 770;
  Best Local Similarity   99.3%;  Pred. No. 1e-114;
  Matches  269;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        107 RCLWEESRLLQTAATAAQQGGQANHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVE 166

Qy         61 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 NLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLS 226

Qy        121 AMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 180
              |||||||||||||||||| :||||||||||||||||||||||||||||||||||||||||
Db        227 AMEYVQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEE 286

Qy        181 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 LQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFT 346

Qy        241 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271
              |||||||||||||||||||||||||||||||
Db        347 TKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


Search completed: May  5, 2005, 15:01:10
Job time : 85.1694 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:14:45 ; Search time 50.566 Seconds
(without alignments)
726.619 Million cell updates/sec

Title: US-10-090-185-21
Perfect score: 477
Sequence: 1 VQDLEQKMKVVENLQDDFDF.....YVQKTLTDEELADWKRRPEI 95

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	477	100.0	95	4	AAY72853	Aay72853 Mouse Sta
2	477	100.0	128	4	AAY72852	Aay72852 Mouse Sta
3	477	100.0	143	4	AAY72849	Aay72849 Mouse Sta
4	477	100.0	176	4	AAY72848	Aay72848 Mouse Sta
5	477	100.0	213	4	AAY72851	Aay72851 Mouse Sta
6	477	100.0	223	4	AAY72854	Aay72854 Mouse Sta
7	477	100.0	228	4	AAY72861	Aay72861 Mouse Sta
8	477	100.0	229	4	AAY72860	Aay72860 Mouse Sta
9	477	100.0	229	4	AAY72850	Aay72850 Mouse Sta
10	477	100.0	229	4	AAY72863	Aay72863 Mouse Sta
11	477	100.0	229	4	AAY72862	Aay72862 Mouse Sta
12	477	100.0	236	4	AAY72847	Aay72847 Mouse Sta
13	477	100.0	252	4	AAY72846	Aay72846 Mouse Sta
14	477	100.0	271	4	AAY72841	Aay72841 Mouse Sta
15	477	100.0	770	2	AAR72082	Aar72082 Mouse Sta
16	477	100.0	770	2	AAW03176	Aaw03176 Mouse STA
17	466	97.7	720	5	AAE22055	Aae22055 Human Sta
18	466	97.7	769	5	ABB57164	Abb57164 Mouse isc
19	466	97.7	769	5	AAE22054	Aae22054 Human Sta
20	466	97.7	769	5	AAE22056	Aae22056 Human pro
21	466	97.7	770	2	AAR82995	Aar82995 Mouse liv
22	466	97.7	770	2	AAR82993	Aar82993 Human pla
23	466	97.7	770	2	AAY03768	Aay03768 Human STA
24	466	97.7	770	3	AAB12377	Aab12377 N-termina
25	466	97.7	770	4	AAB19964	Aab19964 Human sig
26	466	97.7	770	5	AAE14652	Aae14652 Murine ST
27	466	97.7	770	5	ABG69497	Abg69497 Human bai
28	466	97.7	770	5	AAE15174	Aae15174 Human Sta
29	466	97.7	770	6	ABU10476	Abu10476 Mouse STA
30	466	97.7	770	7	ADD44738	Add44738 Rat Prote
31	466	97.7	770	7	ADD44740	Add44740 Human Pro
32	466	97.7	770	8	ADN04365	Adn04365 Antipsori
33	466	97.7	770	8	ADP54789	Adp54789 Human PRO
34	466	97.7	793	3	AAB58442	Aab58442 Lung canc
35	278	58.3	185	4	AAY72855	Aay72855 Mouse Sta
36	173	36.3	423	6	ABR47598	Abr47598 Breast ca
37	173	36.3	704	2	AAW01102	Aaw01102 Signal tr
38	173	36.3	748	2	AAW01101	Aaw01101 Signal tr
39	173	36.3	748	4	AAB19965	Aab19965 Human sig
40	173	36.3	748	5	AAE15175	Aae15175 Human Sta
41	173	36.3	748	6	ABR47599	Abr47599 Breast ca
42	173	36.3	748	8	ADL82891	Adl82891 Human PRO
43	173	36.3	748	8	ADP13003	Adp13003 Protein e
44	173	36.3	748	8	ADR14069	Adr14069 Human NF-
45	173	36.3	748	8	ADR14361	Adr14361 Human NF-

ALIGNMENTS

RESULT 1
AAY72853
ID AAY72853 standard; protein; 95 AA.
XX
AC AAY72853;
XX
DT 31-MAY-2001 (first entry)
XX
DE Mouse Stat3 protein fragment #11 (155-249 amino acids).
XX
KW Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW cellular transformation; dysproliferative disease; cancer; psoriasis;
KW therapy.
XX
OS Mus musculus.
XX
PN WO200116605-A2.
XX
PD 08-MAR-2001.
XX
PF 30-AUG-2000; 2000WO-US023822.
XX
PR 31-AUG-1999; 99US-00387418.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI Zhang X, Horvath C, Wrzeszcynska MH, Darnell JE;
XX
DR WPI; 2001-226705/23.
XX
PT Identifying an agent for use in modulating the interaction between
PT transcription factor c-Jun and a Stat3 protein.
XX
PS Claim 65; Page 78; 86pp; English.
XX
CC The present sequence is mouse Stat3 protein fragment containing 155-249
CC amino acids of Stat3 protein. The invention relates to methods for
CC identifying interacting regions of transcription factors and methods for
CC identifying agents which modulates the interaction between a
CC transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC Stat-3, useful for modulating gene transcription e.g., cellular
CC transformation. These identifying agents are used in the treatment of
CC dysproliferative diseases and also for treating cancer and psoriasis. A
CC Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ Sequence 95 AA;

```
  Query Match           100.0%;  Score 477;  DB 4;  Length 95;
  Best Local Similarity  100.0%;  Pred. No. 5.9e-41;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95


RESULT 2
AAY72852
ID   AAY72852 standard; protein; 128 AA.
XX
AC   AAY72852;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 protein fragment #10 (155-282 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 78; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 155-282
CC   amino acids of Stat3 protein. The invention relates to methods for
CC   identifying interacting regions of transcription factors and methods for
CC   identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 128 AA;

  Query Match           100.0%;  Score 477;  DB 4;  Length 128;
  Best Local Similarity  100.0%;  Pred. No. 8.3e-41;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95


RESULT 3
AAY72849
ID   AAY72849 standard; protein; 143 AA.
XX
AC   AAY72849;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 protein fragment #7 (107-249 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          24. .48
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 75-76; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 107-249
CC   amino acids of Stat3 protein. The invention relates to methods for
CC   identifying interacting regions of transcription factors and methods for
CC   identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 143 AA;

  Query Match           100.0%;  Score 477;  DB 4;  Length 143;
  Best Local Similarity  100.0%;  Pred. No. 9.5e-41;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         49 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 108

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db        109 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143


RESULT 4
AAY72848
ID   AAY72848 standard; protein; 176 AA.
XX
AC   AAY72848;
XX
```

```
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 protein fragment #6 (107-282 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          24. .48
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 74-75; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 107-282
CC   amino acids of Stat3 protein. The invention relates to methods for
CC   identifying interacting regions of transcription factors and methods for
CC   identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 176 AA;

  Query Match             100.0%;  Score 477;  DB 4;  Length 176;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-40;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         49 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 108

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db        109 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143


RESULT 5
AAY72851
ID   AAY72851 standard; protein; 213 AA.
XX
AC   AAY72851;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 protein fragment #9 (130-342 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .25
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 77; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 130-342
CC   amino acids of Stat3 protein. This Stat3 fragment showed very weak
CC   binding to c-Jun protein in the cell extract. The invention relates to
CC   methods for identifying interacting regions of transcription factors and
CC   methods for identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 213 AA;

  Query Match             100.0%;  Score 477;  DB 4;  Length 213;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-40;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 85

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db         86 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120


RESULT 6
AAY72854
ID   AAY72854 standard; protein; 223 AA.
XX
AC   AAY72854;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 protein fragment #12 (155-377 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          188. .204
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
```

```
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 79; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 155-377
CC   amino acids of Stat3 protein. The invention relates to methods for
CC   identifying interacting regions of transcription factors and methods for
CC   identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 223 AA;

  Query Match             100.0%;  Score 477;  DB 4;  Length 223;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-40;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95


RESULT 7
AAY72861
ID   AAY72861 standard; protein; 228 AA.
XX
AC   AAY72861;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 mutant (T346A, K348A, R350A) protein fragment.
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy; mutant; mutein.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .25
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
FT   Region          213. .229
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
FT   Misc-difference 217
FT                   /note= "Wild type Thr substituted with Ala; corresponds
FT                   to 346 position of Stat-3 protein"
FT   Misc-difference 219
FT                   /note= "Wild type Lys substituted with Ala; corresponds
FT                   to 348 position of Stat-3 protein"
FT   Misc-difference 221
FT                   /note= "Wild type Arg substituted with Ala; corresponds
FT                   to 350 position of Stat-3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 66; Page 84-85; 86pp; English.
XX
CC   The present sequence is mouse Stat3 mutant protein fragment containing
CC   130-358 amino acids of Stat3 protein. This mutant is obtained by
CC   replacing Thr 346 with Ala, Lys 348 with Ala and Arg 350 with Ala in the
CC   Stat3 protein. The invention relates to methods for identifying
CC   interacting regions of transcription factors and methods for identifying
CC   agents which modulates the interaction between a transcription factor
CC   such as c-Jun and a Stat protein such as Stat-1 and Stat-3, useful for
CC   modulating gene transcription e.g., cellular transformation. These
CC   identifying agents are used in the treatment of dysproliferative diseases
CC   and also for treating cancer and psoriasis. A Stat protein comprises the
CC   N-terminal domain, coiled-coil domain, DNA binding domain, linker domain,
CC   SH2 domain and transactivation domain
XX
SQ   Sequence 228 AA;

  Query Match             100.0%;  Score 477;  DB 4;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-40;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         25 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 84

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db         85 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 119


RESULT 8
AAY72860
ID   AAY72860 standard; protein; 229 AA.
XX
AC   AAY72860;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 mutant (L148A,V151A,T346A,K348A,R350A) protein fragment.
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy; mutant; mutein.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .25
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
```

```
FT                   130-154 position of Stat3 protein"
FT   Misc-difference 19
FT                   /note= "Wild type Leu substituted with Ala; corresponds
FT                   to 148 position of Stat-3 protein"
FT   Misc-difference 22
FT                   /note= "Wild type Val substituted with Ala corresponds to
FT                   151 position of Stat-3 protein"
FT   Region          213. .229
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
FT   Misc-difference 217
FT                   /note= "Wild type Thr substituted with Ala corresponds to
FT                   346 position of Stat-3 protein"
FT   Misc-difference 219
FT                   /note= "Wild type Lys substituted with Ala corresponds to
FT                   348 position of Stat-3 protein"
FT   Misc-difference 221
FT                   /note= "Wild type Arg substituted with Ala corresponds to
FT                   350 position of Stat-3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Example 4; Page; 86pp; English.
XX
CC   The present sequence is mouse Stat3 mutant protein fragment containing
CC   130-358 amino acids of Stat3 protein. This mutant is obtained by
CC   replacing Leu 148 with Ala, Val 151 with Ala, Thr 346 with Ala, Lys 348
CC   with Ala and Arg 350 with Ala in the Stat3 protein. The invention relates
CC   to methods for identifying interacting regions of transcription factors
CC   and methods for identifying agents which modulates the interaction
CC   between a transcription factor such as c-Jun and a Stat protein such as
CC   Stat-1 and Stat-3, useful for modulating gene transcription e.g.,
CC   cellular transformation. These identifying agents are used in the
CC   treatment of dysproliferative diseases and also for treating cancer and
CC   psoriasis. A Stat protein comprises the N-terminal domain, coiled-coil
CC   domain, DNA binding domain, linker domain, SH2 domain and transactivation
CC   domain
XX
SQ   Sequence 229 AA;

  Query Match             100.0%;  Score 477;  DB 4;  Length 229;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-40;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 85

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db         86 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120


RESULT 9
AAY72850
ID   AAY72850 standard; protein; 229 AA.
XX
AC   AAY72850;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 protein fragment #8 (130-358 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .25
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
FT   Region          213. .229
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 76-77; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 130-358
CC   amino acids of Stat3 protein. This Stat3 fragment showed strong binding
CC   to c-Jun protein in the cell extract. The invention relates to methods
CC   for identifying interacting regions of transcription factors and methods
CC   for identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 229 AA;

  Query Match             100.0%;  Score 477;  DB 4;  Length 229;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-40;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 85

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db         86 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120


RESULT 10
AAY72863
ID   AAY72863 standard; protein; 229 AA.
XX
AC   AAY72863;
XX
DT   31-MAY-2001  (first entry)
XX
```

```
DE   Mouse Stat3 mutant (V151A) protein fragment (130-358 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy; mutant; mutein.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .25
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
FT   Misc-difference 22
FT                   /note= "Wild type Val substituted with Ala corresponds to
FT                   151 position of Stat-3 protein"
FT   Region          213. .229
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 66; Page 86; 86pp; English.
XX
CC   The present sequence is mouse Stat3 mutant (V151A) protein fragment
CC   containing 130-358 amino acids of Stat3 protein. This mutant is obtained
CC   by replacing Val 151 with Ala in Stat3 protein. The invention relates to
CC   methods for identifying interacting regions of transcription factors and
CC   methods for identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 229 AA;

  Query Match             100.0%;  Score 477;  DB 4;  Length 229;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-40;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 85

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db         86 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120


RESULT 11
AAY72862
ID   AAY72862 standard; protein; 229 AA.
XX
AC   AAY72862;
XX
DT   31-MAY-2001  (first entry)
XX
```

```
DE   Mouse Stat3 mutant (L148A) protein fragment (130-358 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy; mutant; mutein.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .25
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
FT   Misc-difference 19
FT                   /note= "Wild type Leu substituted with Ala; corresponds
FT                   to 148 position of Stat-3 protein"
FT   Region          213. .229
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 66; Page 85; 86pp; English.
XX
CC   The present sequence is mouse Stat3 mutant (L148A) protein fragment
CC   containing 130-358 amino acids of Stat3 protein. This mutant is obtained
CC   by replacing Leu 148 with Ala in Stat3 protein. The invention relates to
CC   methods for identifying interacting regions of transcription factors and
CC   methods for identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 229 AA;

  Query Match             100.0%;  Score 477;  DB 4;  Length 229;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-40;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 85

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db         86 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120


RESULT 12
AAY72847
ID   AAY72847 standard; protein; 236 AA.
XX
AC   AAY72847;
XX
DT   31-MAY-2001  (first entry)
XX
```

```
DE   Mouse Stat3 protein fragment #5 (107-342 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          24. .48
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 74; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 107-342
CC   amino acids of Stat3 protein. The invention relates to methods for
CC   identifying interacting regions of transcription factors and methods for
CC   identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 236 AA;

  Query Match             100.0%;  Score 477;  DB 4;  Length 236;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-40;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         49 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 108

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db        109 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143


RESULT 13
AAY72846
ID   AAY72846 standard; protein; 252 AA.
XX
AC   AAY72846;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 protein fragment #4 (107-358 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          24. .48
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
FT   Region          236. .252
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 73; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 107-358
CC   amino acids of Stat3 protein. This Stat3 fragment showed strong binding
CC   to c-Jun protein in the cell extract. The invention relates to methods
CC   for identifying interacting regions of transcription factors and methods
CC   for identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 252 AA;

  Query Match             100.0%;  Score 477;  DB 4;  Length 252;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-40;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         49 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 108

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db        109 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143


RESULT 14
AAY72841
ID   AAY72841 standard; protein; 271 AA.
XX
AC   AAY72841;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 protein fragment #2 (107-377 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          24. .48
```

```
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
FT   Region          236. .252
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 67-68; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 107-377
CC   amino acids of Stat3 protein. This Stat3 fragment showed strong binding
CC   to c-Jun protein in the cell extract. The invention relates to methods
CC   for identifying interacting regions of transcription factors and methods
CC   for identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 271 AA;

  Query Match             100.0%;  Score 477;  DB 4;  Length 271;
  Best Local Similarity   100.0%;  Pred. No. 2e-40;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         49 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 108

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db        109 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143


RESULT 15
AAR72082
ID   AAR72082 standard; protein; 770 AA.
XX
AC   AAR72082;
XX
DT   25-MAR-2003  (revised)
DT   27-SEP-1995  (first entry)
XX
DE   Mouse Stat3 (19sf6).
XX
KW   Signal transducer and activator of transcription; STAT; 19sf6; Stat3;
KW   receptor recognition factor; transcription factor; cellular debilitation;
KW   derangement; dysfunction; interferon-gamma.
XX
OS   Mus sp.
XX
PN   WO9508629-A1.
XX
PD   30-MAR-1995.
XX
PF   26-SEP-1994;   94WO-US010849.
XX
PR   24-SEP-1993;   93US-00126588.
PR   24-SEP-1993;   93US-00126595.
PR   11-MAR-1994;   94US-00212184.
PR   11-MAR-1994;   94US-00212185.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Darnell JE,  Schindler CW,  Shuai K,  Wen Z,  Zhong Z;
XX
DR   WPI; 1995-139598/18.
DR   N-PSDB; AAQ89340.
XX
PT   Receptor recognition factor implicated in transcriptional stimulation of
PT   genes - useful in drug screening assays and/or for treating cellular
PT   debilitations, derangements and/or dysfunctions, etc.
XX
PS   Claim 1; Page 107-110; 160pp; English.
XX
CC   A fragment encoding the human Stat91 protein was used to screen a murine
CC   thymus and spleen cDNA for homologous proteins. A highly homologous gene
CC   (given in AAQ89338) was isolated that encoded a 91 kDa protein (AAR72080)
CC   (Stat1) that was responsive to interferon- gamma. Using a fragment of the
CC   mouse gene as probe, 2 additional members of the 113-91 family of
CC   receptor recognition factor proteins were isolated. The 2 genes (AAQ89339
CC   -40) were cloned in plasmids 13sf1 and 19sf6 and encoded proteins termed
CC   Stat4 (AAR72081) and Stat3 (AAR72082), respectively. (Updated on 25-MAR-
CC   2003 to correct PN field.)
XX
SQ   Sequence 770 AA;

  Query Match             100.0%;  Score 477;  DB 2;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 6.8e-40;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 214

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db        215 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249


Search completed: May 25, 2005, 17:36:34
Job time : 58.7327 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:25:01 ; Search time 11.8868 Seconds
(without alignments)
596.600 Million cell updates/sec

Title: US-10-090-185-21
Perfect score: 477
Sequence: 1 VQDLEQKMKVVENLQDDFDF.....YVQKTLTDEELADWKRRPEI 95

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	477	100.0	95	3	US-09-387-418A-21	Sequence 21, Appl
2	477	100.0	128	3	US-09-387-418A-20	Sequence 20, Appl
3	477	100.0	143	3	US-09-387-418A-17	Sequence 17, Appl
4	477	100.0	176	3	US-09-387-418A-16	Sequence 16, Appl
5	477	100.0	213	3	US-09-387-418A-19	Sequence 19, Appl
6	477	100.0	223	3	US-09-387-418A-22	Sequence 22, Appl
7	477	100.0	229	3	US-09-387-418A-18	Sequence 18, Appl
8	477	100.0	229	3	US-09-387-418A-28	Sequence 28, Appl
9	477	100.0	229	3	US-09-387-418A-29	Sequence 29, Appl
10	477	100.0	229	3	US-09-387-418A-30	Sequence 30, Appl
11	477	100.0	229	3	US-09-387-418A-31	Sequence 31, Appl
12	477	100.0	236	3	US-09-387-418A-15	Sequence 15, Appl
13	477	100.0	252	3	US-09-387-418A-14	Sequence 14, Appl
14	477	100.0	271	3	US-09-387-418A-9	Sequence 9, Appli
15	477	100.0	770	1	US-08-369-796-12	Sequence 12, Appl
16	477	100.0	770	2	US-08-852-091-12	Sequence 12, Appl
17	477	100.0	770	2	US-08-820-754-12	Sequence 12, Appl
18	477	100.0	770	3	US-08-956-652-12	Sequence 12, Appl
19	477	100.0	770	3	US-08-956-869-12	Sequence 12, Appl
20	477	100.0	770	3	US-08-948-547-12	Sequence 12, Appl
21	477	100.0	770	3	US-09-364-970-3	Sequence 3, Appli
22	477	100.0	770	3	US-09-364-970-5	Sequence 5, Appli
23	477	100.0	770	3	US-08-956-653A-12	Sequence 12, Appl
24	477	100.0	770	4	US-08-212-185-12	Sequence 12, Appl
25	477	100.0	770	5	PCT-US95-17025-12	Sequence 12, Appl
26	466	97.7	770	1	US-08-416-581B-1	Sequence 1, Appli
27	466	97.7	770	1	US-08-416-581B-5	Sequence 5, Appli
28	466	97.7	770	1	US-08-416-581B-9	Sequence 9, Appli
29	466	97.7	770	3	US-09-012-710-8	Sequence 8, Appli
30	466	97.7	770	3	US-09-087-465-6	Sequence 6, Appli
31	466	97.7	770	3	US-09-556-273-8	Sequence 8, Appli
32	466	97.7	770	3	US-09-526-542-2	Sequence 2, Appli
33	466	97.7	770	4	US-09-972-800A-6	Sequence 6, Appli
34	466	97.7	770	4	US-10-117-087-2	Sequence 2, Appli
35	466	97.7	771	1	US-08-276-099A-14	Sequence 14, Appl
36	466	97.7	771	1	US-08-781-890-14	Sequence 14, Appl
37	278	58.3	185	3	US-09-387-418A-23	Sequence 23, Appl
38	173	36.3	748	1	US-08-408-318-2	Sequence 2, Appli
39	173	36.3	748	1	US-08-839-164-2	Sequence 2, Appli
40	173	36.3	748	3	US-09-087-465-8	Sequence 8, Appli
41	173	36.3	748	4	US-09-972-800A-8	Sequence 8, Appli
42	170	35.6	749	1	US-08-369-796-8	Sequence 8, Appli
43	170	35.6	749	2	US-08-852-091-8	Sequence 8, Appli
44	170	35.6	749	2	US-08-820-754-8	Sequence 8, Appli
45	170	35.6	749	3	US-08-956-652-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-387-418A-21
; Sequence 21, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/09/387,418A
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-387-418A-21

Query Match 100.0%; Score 477; DB 3; Length 95;
Best Local Similarity 100.0%; Pred. No. 3.8e-45;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
```

RESULT 2
US-09-387-418A-20
; Sequence 20, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/09/387,418A
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43

```
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 20
;   LENGTH: 128
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-20

  Query Match             100.0%;  Score 477;  DB 3;  Length 128;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-45;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95


RESULT 3
US-09-387-418A-17
; Sequence 17, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
;  APPLICANT:  Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 17
;   LENGTH: 143
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-17

  Query Match             100.0%;  Score 477;  DB 3;  Length 143;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-45;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         49 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 108

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db        109 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143


RESULT 4
US-09-387-418A-16
; Sequence 16, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
;  APPLICANT:  Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 16
;   LENGTH: 176
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-16

  Query Match             100.0%;  Score 477;  DB 3;  Length 176;
  Best Local Similarity   100.0%;  Pred. No. 8.2e-45;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         49 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 108

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db        109 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143


RESULT 5
US-09-387-418A-19
; Sequence 19, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
;  APPLICANT:  Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 19
;   LENGTH: 213
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-19

  Query Match             100.0%;  Score 477;  DB 3;  Length 213;
  Best Local Similarity   100.0%;  Pred. No. 1e-44;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 85

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db         86 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120


RESULT 6
US-09-387-418A-22
; Sequence 22, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
;  APPLICANT:  Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 22
;   LENGTH: 223
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-22
```

Query Match 100.0%; Score 477; DB 3; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
```

RESULT 7
US-09-387-418A-18
; Sequence 18, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/09/387,418A
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-387-418A-18

Query Match 100.0%; Score 477; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 85

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db         86 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
```

RESULT 8
US-09-387-418A-28
; Sequence 28, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/09/387,418A
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-387-418A-28

Query Match 100.0%; Score 477; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 85

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db         86 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
```

RESULT 9
US-09-387-418A-29
; Sequence 29, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/09/387,418A
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-387-418A-29

Query Match 100.0%; Score 477; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 85

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db         86 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
```

RESULT 10
US-09-387-418A-30
; Sequence 30, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/09/387,418A
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-387-418A-30

Query Match 100.0%; Score 477; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 1.1e-44;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 85

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db         86 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120


RESULT 11
US-09-387-418A-31
; Sequence 31, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
;  APPLICANT:  Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 31
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-31

  Query Match             100.0%;  Score 477;  DB 3;  Length 229;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-44;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 85

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db         86 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120


RESULT 12
US-09-387-418A-15
; Sequence 15, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
;  APPLICANT:  Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 15
;   LENGTH: 236
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-15

  Query Match             100.0%;  Score 477;  DB 3;  Length 236;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-44;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         49 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 108

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db        109 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143


RESULT 13
US-09-387-418A-14
; Sequence 14, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
;  APPLICANT:  Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 252
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-14

  Query Match             100.0%;  Score 477;  DB 3;  Length 252;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-44;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         49 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 108

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db        109 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143


RESULT 14
US-09-387-418A-9
; Sequence 9, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
;  APPLICANT:  Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 271
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-9

  Query Match             100.0%;  Score 477;  DB 3;  Length 271;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-44;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         49 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 108

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db        109 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143
```

```
RESULT 15
US-08-369-796-12
; Sequence 12, Application US/08369796
; Patent No. 5716622
;  GENERAL INFORMATION:
;    APPLICANT:  James E. Darnell, Jr.
;    APPLICANT:  Zilong Wen
;    APPLICANT:  Curt M. Horvath
;    APPLICANT:  Zhong Zhong
;    TITLE OF INVENTION:  FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
;    TITLE OF INVENTION:  TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/369,796
;      FILING DATE:  06-JAN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-116
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-369-796-12

  Query Match             100.0%;  Score 477;  DB 1;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-44;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 214

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db        215 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249


Search completed: May 25, 2005, 17:47:37
Job time : 15.8868 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:36:58 ; Search time 41.6038 Seconds
(without alignments)
763.830 Million cell updates/sec

Title: US-10-090-185-21
Perfect score: 477
Sequence: 1 VQDLEQKMKVVENLQDDFDF.....YVQKTLTDEELADWKRRPEI 95

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1434725 seqs, 334507595 residues

Total number of hits satisfying chosen parameters: 1434725

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	477	100.0	95	13	US-10-090-185-21	Sequence 21, Appl
2	477	100.0	128	13	US-10-090-185-20	Sequence 20, Appl
3	477	100.0	143	13	US-10-090-185-17	Sequence 17, Appl
4	477	100.0	176	13	US-10-090-185-16	Sequence 16, Appl
5	477	100.0	213	13	US-10-090-185-19	Sequence 19, Appl
6	477	100.0	223	13	US-10-090-185-22	Sequence 22, Appl
7	477	100.0	229	13	US-10-090-185-18	Sequence 18, Appl
8	477	100.0	229	13	US-10-090-185-28	Sequence 28, Appl
9	477	100.0	229	13	US-10-090-185-29	Sequence 29, Appl
10	477	100.0	229	13	US-10-090-185-30	Sequence 30, Appl
11	477	100.0	229	13	US-10-090-185-31	Sequence 31, Appl
12	477	100.0	236	13	US-10-090-185-15	Sequence 15, Appl
13	477	100.0	252	13	US-10-090-185-14	Sequence 14, Appl
14	477	100.0	271	13	US-10-090-185-9	Sequence 9, Appli
15	477	100.0	770	11	US-09-876-773-12	Sequence 12, Appl
16	477	100.0	770	17	US-10-639-617-12	Sequence 12, Appl
17	466	97.7	720	15	US-10-380-020-4	Sequence 4, Appli
18	466	97.7	769	15	US-10-380-020-2	Sequence 2, Appli
19	466	97.7	769	15	US-10-380-020-5	Sequence 5, Appli
20	466	97.7	770	14	US-10-045-792-8	Sequence 8, Appli
21	466	97.7	770	14	US-10-038-010-56	Sequence 56, Appl
22	466	97.7	770	14	US-10-117-087-2	Sequence 2, Appli
23	466	97.7	770	15	US-10-116-275-329	Sequence 329, App
24	466	97.7	770	15	US-10-116-275-349	Sequence 349, App
25	466	97.7	793	9	US-09-925-302-780	Sequence 780, App
26	466	97.7	793	10	US-09-925-302-780	Sequence 780, App
27	278	58.3	185	13	US-10-090-185-23	Sequence 23, Appl
28	173	36.3	423	14	US-10-177-293-437	Sequence 437, App
29	173	36.3	748	14	US-10-177-293-439	Sequence 439, App
30	173	36.3	748	16	US-10-755-889-70	Sequence 70, Appl
31	173	36.3	748	16	US-10-755-889-362	Sequence 362, App
32	170	35.6	749	9	US-09-833-205-6	Sequence 6, Appli
33	170	35.6	749	11	US-09-876-773-8	Sequence 8, Appli
34	170	35.6	749	14	US-10-045-792-7	Sequence 7, Appli
35	170	35.6	749	14	US-10-205-194-39	Sequence 39, Appl
36	170	35.6	749	17	US-10-639-617-8	Sequence 8, Appli
37	168	35.2	749	9	US-09-833-205-4	Sequence 4, Appli
38	160.5	33.6	749	14	US-10-045-792-9	Sequence 9, Appli
39	158	33.1	268	13	US-10-090-185-12	Sequence 12, Appl
40	158	33.1	582	14	US-10-245-120-3	Sequence 3, Appli
41	158	33.1	712	11	US-09-876-773-6	Sequence 6, Appli
42	158	33.1	712	14	US-10-245-120-2	Sequence 2, Appli
43	158	33.1	712	17	US-10-639-617-6	Sequence 6, Appli
44	158	33.1	712	17	US-10-936-390-5	Sequence 5, Appli
45	158	33.1	750	9	US-09-833-205-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-10-090-185-21
; Sequence 21, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/10/090,185
; CURRENT FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: 09/387,418
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-090-185-21

Query Match 100.0%; Score 477; DB 13; Length 95;
Best Local Similarity 100.0%; Pred. No. 1.8e-42;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
```

```
RESULT 2
US-10-090-185-20
; Sequence 20, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
;   LENGTH: 128
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-20

  Query Match            100.0%;  Score 477;  DB 13;  Length 128;
  Best Local Similarity  100.0%;  Pred. No. 2.6e-42;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95


RESULT 3
US-10-090-185-17
; Sequence 17, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
;   LENGTH: 143
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-17

  Query Match            100.0%;  Score 477;  DB 13;  Length 143;
  Best Local Similarity  100.0%;  Pred. No. 2.9e-42;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         49 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 108

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db        109 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143


RESULT 4
US-10-090-185-16
; Sequence 16, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
;   LENGTH: 176
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-16

  Query Match            100.0%;  Score 477;  DB 13;  Length 176;
  Best Local Similarity  100.0%;  Pred. No. 3.8e-42;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         49 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 108

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db        109 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143


RESULT 5
US-10-090-185-19
; Sequence 19, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
;   LENGTH: 213
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-19

  Query Match            100.0%;  Score 477;  DB 13;  Length 213;
  Best Local Similarity  100.0%;  Pred. No. 4.8e-42;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 85

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
```

```
Db          86 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120


RESULT 6
US-10-090-185-22
; Sequence 22, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
;   LENGTH: 223
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-22

  Query Match             100.0%;  Score 477;  DB 13;  Length 223;
  Best Local Similarity   100.0%;  Pred. No. 5.1e-42;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60

Qy          61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
               |||||||||||||||||||||||||||||||||||
Db          61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95


RESULT 7
US-10-090-185-18
; Sequence 18, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-18

  Query Match             100.0%;  Score 477;  DB 13;  Length 229;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-42;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          26 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 85

Qy          61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
               ||||||||||||||||||||||||||||||||||||
Db          86 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120


RESULT 8
US-10-090-185-28
; Sequence 28, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-28

  Query Match             100.0%;  Score 477;  DB 13;  Length 229;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-42;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          26 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 85

Qy          61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
               ||||||||||||||||||||||||||||||||||||
Db          86 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120


RESULT 9
US-10-090-185-29
; Sequence 29, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-29

  Query Match             100.0%;  Score 477;  DB 13;  Length 229;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-42;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          26 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 85
```

```
Qy          61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
               |||||||||||||||||||||||||||||||||||||
Db          86 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
```

RESULT 10
US-10-090-185-30
; Sequence 30, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/10/090,185
; CURRENT FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: 09/387,418
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-090-185-30

Query Match 100.0%; Score 477; DB 13; Length 229;
Best Local Similarity 100.0%; Pred. No. 5.3e-42;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy           1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          26 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 85

Qy          61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
               |||||||||||||||||||||||||||||||||||||
Db          86 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
```

RESULT 11
US-10-090-185-31
; Sequence 31, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/10/090,185
; CURRENT FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: 09/387,418
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-090-185-31

Query Match 100.0%; Score 477; DB 13; Length 229;
Best Local Similarity 100.0%; Pred. No. 5.3e-42;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy           1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          26 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 85

Qy          61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
               |||||||||||||||||||||||||||||||||||||
Db          86 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
```

RESULT 12
US-10-090-185-15
; Sequence 15, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/10/090,185
; CURRENT FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: 09/387,418
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-090-185-15

Query Match 100.0%; Score 477; DB 13; Length 236;
Best Local Similarity 100.0%; Pred. No. 5.5e-42;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy           1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          49 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 108

Qy          61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
               |||||||||||||||||||||||||||||||||||||
Db         109 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143
```

RESULT 13
US-10-090-185-14
; Sequence 14, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/10/090,185
; CURRENT FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: 09/387,418
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-090-185-14

Query Match 100.0%; Score 477; DB 13; Length 252;
Best Local Similarity 100.0%; Pred. No. 5.9e-42;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy           1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          49 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 108

Qy          61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
               |||||||||||||||||||||||||||||||||||
Db         109 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143


RESULT 14
US-10-090-185-9
; Sequence 9, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 271
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-9

  Query Match             100.0%;  Score 477;  DB 13;  Length 271;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-42;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          49 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 108

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db        109 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143


RESULT 15
US-09-876-773-12
; Sequence 12, Application US/09876773
; Publication No. US20040058318A1
;   GENERAL INFORMATION:
;        APPLICANT: Darnell Jr., James E.
;                   Schindler, Christian W.
;                   Fu, Xian-Yuan
;                   Wen, Zilong
;                   Zhong, Zhong
;        TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
;                            SEQUENCES AND METHODS OF USE THEREOF
;        NUMBER OF SEQUENCES: 25
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Klauber & Jackson
;             STREET: 411 Hackensack Avenue
;             CITY: Hackensack
;             STATE: New Jersey
;             COUNTRY: USA
;             ZIP: 07601
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/876,773
;             FILING DATE: 07-Jun-2001
;             CLASSIFICATION: 435
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/212,185
;             FILING DATE: 11-MAR-1994
;             APPLICATION NUMBER: US 07/980,498
;             FILING DATE: 23-NOV-1992
;             APPLICATION NUMBER: US 07/854,296
;             FILING DATE: 19-MAR-1992
;             APPLICATION NUMBER: WO US93/02569
;             FILING DATE: 19-MAR-1993
;             APPLICATION NUMBER: US 08/126,588
;             FILING DATE: 24-SEP-1993
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Jackson Esq., David A.
;             REGISTRATION NUMBER: 26,742
;             REFERENCE/DOCKET NUMBER: 600-1-073 CIP
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 201 487-5800
;             TELEFAX: 201 343-1684
;             TELEX: 133521
;   INFORMATION FOR SEQ ID NO: 12:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 770 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-876-773-12

  Query Match             100.0%;  Score 477;  DB 11;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-41;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 214

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||||||
Db        215 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249


Search completed: May 25, 2005, 18:21:48
Job time : 53.9371 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:24:07 ; Search time 8.96226 Seconds
(without alignments)
1019.898 Million cell updates/sec

Title: US-10-090-185-21
Perfect score: 477
Sequence: 1 VQDLEQKMKVVENLQDDFDF.....YVQKTLTDEELADWKRRPEI 95

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	466	97.7	770	2	A54444	DNA-binding protei
2	466	97.7	770	2	I49508	ISGF3 p91-related
3	164	34.4	748	2	A56047	gamma-interferon a
4	158	33.1	739	2	A46159	interferon-depende
5	93	19.5	786	2	I49274	mammary gland fact
6	93	19.5	793	2	S54772	mammary gland fact
7	89	18.7	794	2	G02317	transcription acti
8	88	18.4	794	2	S55527	mammary gland fact
9	86	18.0	533	2	G72593	hypothetical prote
10	85.5	17.9	2094	2	S33124	tpr protein - huma
11	84.5	17.7	217	2	G75219	hypothetical prote
12	84.5	17.7	821	2	T24728	hypothetical prote
13	83.5	17.5	3187	2	JC5837	364K Golgi complex
14	82	17.2	1818	1	S73852	hypothetical prote
15	81.5	17.1	217	2	B71203	hypothetical prote
16	79.5	16.7	278	2	AC2354	hypothetical prote
17	79	16.6	638	2	S67605	hypothetical prote
18	78	16.4	1968	1	S05697	myosin heavy chain
19	77.5	16.2	1160	2	I40589	parasporal crystal
20	77	16.1	304	2	AD1029	probable membrane
21	77	16.1	569	2	B71902	hypothetical prote
22	76.5	16.0	441	2	T01613	hypothetical prote
23	76	15.9	821	2	S67087	hypothetical prote
24	75.5	15.8	734	2	T27055	hypothetical prote
25	75.5	15.8	1085	2	F96712	hypothetical prote
26	75	15.7	742	1	S58691	kinesin-related pr
27	75	15.7	750	2	T38435	coiled coil protei
28	75	15.7	896	2	S43074	epidermal growth f
29	75	15.7	1171	2	T45706	chromosome-associa
30	74.5	15.6	1979	1	S03166	myosin heavy chain
31	74	15.5	412	2	S07537	myosin heavy chain
32	74	15.5	621	2	S10450	myosin heavy chain
33	73.5	15.4	398	2	AC1763	peptidoglycan lyti
34	73.5	15.4	927	2	AG1739	transmembrane prot
35	73.5	15.4	980	2	E71606	hypothetical prote
36	73	15.3	764	2	I51302	myosin heavy chain
37	73	15.3	1526	2	T41522	myosin ii - fissio
38	73	15.3	1738	2	T14867	interaptin - slime
39	72.5	15.2	986	2	T10754	cis-Golgi matrix p
40	72.5	15.2	1046	2	T42734	cytoplasmic linker
41	72.5	15.2	1133	2	T22976	hypothetical prote
42	72.5	15.2	2024	2	A54103	centrosome autoant
43	72	15.1	173	2	JC5610	troponin I - sea s
44	72	15.1	262	2	JC1237	apolipoprotein A-I
45	72	15.1	264	2	S22420	apolipoprotein A-I

ALIGNMENTS

RESULT 1
A54444
DNA-binding protein APRF - human
C;Species: Homo sapiens (man)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 09-Jul-2004
C;Accession: A54444
R;Akira, S.; Nishio, Y.; Inoue, M.; Wang, X.J.; Wei, S.; Matsusaka, T.; Yoshida, K.; Su
Cell 77, 63-71, 1994
A;Title: Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related tr
A;Reference number: A54444; MUID:94208062; PMID:7512451
A;Accession: A54444
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-770 <RES>
A;Cross-references: UNIPROT:P40763; GB:L29277; NID:g475788; PID:g475789
C;Genetics:
A;Gene: GDB:STAT3; APRF
A;Cross-references: GDB:358950
A;Map position: 17q21-17q21
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: DNA binding; transcription factor

Query Match 97.7%; Score 466; DB 2; Length 770;
Best Local Similarity 97.9%; Pred. No. 1.3e-31;
Matches 93; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 214

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              ||||||||||||||||||||||||||||||||| :|
Db        215 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249
```

RESULT 2
I49508
ISGF3 p91-related transcription factor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I49508; I49009
R;Akira, S.; Nishio, Y.; Inoue, M.; Wang, X.J.; Wei, S.; Matsusaka, T.; Yoshida, K.; Su
Cell 77, 63-71, 1994
A;Title: Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related tr
A;Reference number: A54444; MUID:94208062; PMID:7512451
A;Accession: I49508
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-770 <RES>
A;Cross-references: UNIPROT:P42227; GB:L29278; NID:g476715; PIDN:AAA37254.1; PID:g47671
R;Raz, R.; Durbin, J.E.; Levy, D.E.
J. Biol. Chem. 269, 24391-24395, 1994

```
A;Title: Acute phase response factor and additional members of the interferon-stimulated
A;Reference number: I49009; MUID:95014185; PMID:7523373
A;Accession: I49009
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-393,'M',395-700,702-770 <RE2>
A;Cross-references: EMBL:U08378; NID:g473889; PIDN:AAA56668.1; PID:g473890
C;Genetics:
A;Gene: APRF
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             97.7%;  Score 466;  DB 2;  Length 770;
  Best Local Similarity   97.9%;  Pred. No. 1.3e-31;
  Matches   93;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 214

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              ||||||||||||||||||||||||||||||||| :|
Db        215 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249


RESULT 3
A56047
gamma-interferon activation site-binding protein Stat4 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004
C;Accession: A56047
R;Yamamoto, K.; Quelle, F.W.; Thierfelder, W.E.; Kreider, B.L.; Gilbert, D.J.; Jenkins,
Mol. Cell. Biol. 14, 4342-4349, 1994
A;Title: Stat4, a novel gamma interferon activation site-binding protein expressed in ea
A;Reference number: A56047; MUID:94277038; PMID:8007943
A;Accession: A56047
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-748 <YAM>
A;Cross-references: UNIPROT:P42228; GB:U09351; NID:g509502; PIDN:AAA19692.1; PID:g509503
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: DNA binding; phosphoprotein

  Query Match             34.4%;  Score 164;  DB 2;  Length 748;
  Best Local Similarity   35.1%;  Pred. No. 2.8e-06;
  Matches   34;  Conservative   26;  Mismatches   29;  Indels    8;  Gaps    2;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKS--QGDMQDLNGNNQSVTRQKMQQLEQMLTALDQ 58
              ||  ||  | :|:|||:||: |||:::  |||       |  :  |::  |::|| :||
Db        154 VQMTEQDTKYLEDLQDEFDYRYKTIQTMDQGD------KNSILVNQEVLTLQEMLNSLDF 207

Qy         59 MRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |:  :|::  :::  : :  ::  ||| |||:| :|
Db        208 KRKEALSKMTQIVNETDLLMNSMLLEELQDWKKRQQI 244


RESULT 4
A46159
interferon-dependent positive-acting transcription factor ISGF-3 91K chain - human
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A46159
R;Schindler, C.; Fu, X.Y.; Improta, T.; Aebersold, R.; Darnell Jr., J.E.
Proc. Natl. Acad. Sci. U.S.A. 89, 7836-7839, 1992
A;Title: Proteins of transcription factor ISGF-3: one gene encodes the 91-and 84-kDa ISG
A;Reference number: A46159; MUID:92366557; PMID:1502203
A;Accession: A46159
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid; protein
A;Residues: 1-739 <SCH>
A;Cross-references: UNIPROT:P42224
A;Experimental source: HeLa cells
A;Note: sequence extracted from NCBI backbone (NCBIP:110818)
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             33.1%;  Score 158;  DB 2;  Length 739;
  Best Local Similarity   36.2%;  Pred. No. 8.9e-06;
  Matches   34;  Conservative   22;  Mismatches   36;  Indels    2;  Gaps    1;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              |  :| ::| :|:|||::||  |||  |    : ||  :|  :|:  |::|   ||  |
Db        153 VMCIEHEIKSLEDLQDEYDFKCKTL--QNREHETNGVAKSDQKQEQLLLKKMYLMLDNKR 210

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPE 94
              : :| ::  ||:  |  |  | ::|| :|||| :
Db        211 KEVVHKIIELLNVTELTQNALINDELVEWKRRQQ 244


RESULT 5
I49274
mammary gland factor - mouse
N;Alternate names: STAT5 protein homolog p80
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I49274; S54773; S54727
R;Liu, X.; Robinson, G.W.; Gouilleux, F.; Groner, B.; Hennighausen, L.
Proc. Natl. Acad. Sci. U.S.A. 92, 8831-8835, 1995
A;Title: Cloning and expression of Stat5 and an additional homologue (Stat5b) involved
A;Reference number: I49273; MUID:96004632; PMID:7568026
A;Accession: I49274
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-786 <RES>
A;Cross-references: UNIPROT:P42232; UNIPROT:Q9JKM1; EMBL:U21110; NID:g747973; PIDN:AAC5
R;Mui, A.L.F.; Wakao, H.; O'Farrell, A.M.; Harada, N.; Miyajima, A.
EMBO J. 14, 1166-1175, 1995
A;Title: Interleukin-3, granulocyte-macrophage colony stimulating factor and interleuki
A;Reference number: S54772; MUID:95237198; PMID:7720707
A;Accession: S54773
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-432,'E',434-786 <MUI>
A;Cross-references: EMBL:Z48539; NID:g758635; PIDN:CAA88420.1; PID:g758636
R;Azam, M.; Erdjument-Bromage, H.; Kreider, B.L.; Xia, M.; Quelle, F.; Basu, R.; Saris,
EMBO J. 14, 1402-1411, 1995
A;Title: Interleukin-3 signals through multiple isoforms of Stat5.
A;Reference number: S54725; MUID:95246733; PMID:7537213
A;Accession: S54727
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-432,'E',434-786 <AZA>
C;Genetics:
A;Gene: Stat5b
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             19.5%;  Score 93;  DB 2;  Length 786;
  Best Local Similarity   26.4%;  Pred. No. 2.7;
  Matches   28;  Conservative   18;  Mismatches   48;  Indels   12;  Gaps    3;

Qy          2 QDLEQKMKVVENLQDDFDFNY-KTLKSQGDMQDLNGNN-------QSVTRQKMQQLEQML 53
              || | ::| ::  |: |   | ::|: |     |   |        ::  :||   ||  |
Db        156 QDTENELKKLQQTQEYFIIQYQESLRIQAQFAQLGQLNPQERMSRETALQQKQVSLETWL 215

Qy         54 ----TALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
                  | | |  :  :    |  :   |  : |:||  |||| ::
Db        216 QREAQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELIQWKRRQQL 261


RESULT 6
S54772
mammary gland factor - mouse
N;Alternate names: stat5 protein
C;Species: Mus musculus (house mouse)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S54772; I49273
R;Mui, A.L.F.; Wakao, H.; O'Farrell, A.M.; Harada, N.; Miyajima, A.
EMBO J. 14, 1166-1175, 1995
```

```
A;Title: Interleukin-3, granulocyte-macrophage colony stimulating factor and interleukin
A;Reference number: S54772; MUID:95237198; PMID:7720707
A;Accession: S54772
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-793 <MUI>
A;Cross-references: UNIPROT:P42230; UNIPROT:Q9JIA0; EMBL:Z48538; NID:g758633; PIDN:CAA88
R;Liu, X.; Robinson, G.W.; Gouilleux, F.; Groner, B.; Hennighausen, L.
Proc. Natl. Acad. Sci. U.S.A. 92, 8831-8835, 1995
A;Title: Cloning and expression of Stat5 and an additional homologue (Stat5b) involved i
A;Reference number: I49273; MUID:96004632; PMID:7568026
A;Accession: I49273
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-793 <RES>
A;Cross-references: EMBL:U21103; NID:g747971; PIDN:AAA80590.1; PID:g747972
C;Genetics:
A;Gene: Stat5a
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             19.5%;  Score 93;  DB 2;  Length 793;
  Best Local Similarity   26.4%;  Pred. No. 2.7;
  Matches   28;  Conservative   18;  Mismatches   48;  Indels   12;  Gaps    3;

Qy          2 QDLEQKMKVVENLQDDFDFNY-KTLKSQGDMQDLNGNN-------QSVTRQKMQQLEQML 53
              || | ::| ::  |: |   | ::|: |     |   |        ::  :||  || |
Db        156 QDTENELKKLQQTQEYFIIQYQESLRIQAQFAQLGQLNPQERMSRETALQQKQVSLETWL 215

Qy         54 ----TALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
                  | | |  :  :    |  :   |  : |:||  |||| ::
Db        216 QREAQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELIQWKRRQQL 261


RESULT 7
G02317
transcription activator stat5A - human
C;Species: Homo sapiens (man)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 09-Jul-2004
C;Accession: G02317
R;Lin, J.
submitted to the EMBL Data Library, December 1995
A;Reference number: H01043
A;Accession: G02317
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-794 <LIN>
A;Cross-references: UNIPROT:P42229; EMBL:U43185; NID:g1151169; PIDN:AAB06589.1; PID:g115
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             18.7%;  Score 89;  DB 2;  Length 794;
  Best Local Similarity   25.5%;  Pred. No. 6;
  Matches   27;  Conservative   19;  Mismatches   48;  Indels   12;  Gaps    3;

Qy          2 QDLEQKMKVVENLQDDFDFNY-KTLKSQGDMQDL-------NGNNQSVTRQKMQQLEQML 53
              || | ::| ::  |: |   | ::|: |         |       : ::  :||  || |
Db        156 QDTENELKKLQQTQEYFIIQYQESLRIQAQFAQLAQLSPQERLSRETALQQKQVSLEAWL 215

Qy         54 ----TALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
                  | | |  :  :    |  :   |  : |:||  |||| ::
Db        216 QREAQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELIQWKRRQQL 261


RESULT 8
S55527
mammary gland factor - sheep
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S55527; S44353
R;Wakao, H.; Gouilleux, F.; Groner, B.
EMBO J. 14, 854-855, 1995
A;Title: Corrigenda. Mammary gland factor (MGF) is a novel member of the cytokine regula
A;Reference number: S55527; MUID:95188889; PMID:7882987
A;Accession: S55527
A;Molecule type: mRNA
A;Residues: 1-794 <WAK>
A;Cross-references: UNIPROT:P42231; EMBL:X78428; NID:g602354; PIDN:CAA55191.1; PID:g602
A;Note: this is a revision to the sequence from reference S44353
R;Wakao, H.; Gouilleux, F.; Groner, B.
EMBO J. 13, 2182-2191, 1994
A;Title: Mammary gland factor (MGF) is a novel member of the cytokine regulated transcr
A;Reference number: S44353; MUID:94244619; PMID:7514531
A;Accession: S44353
A;Molecule type: mRNA
A;Residues: 1-716,'RHLHGPGSLPSR',729,'P',731,'ASL' <WAW>
A;Cross-references: EMBL:X78428
A;Note: this sequence has been revised in reference S55527
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             18.4%;  Score 88;  DB 2;  Length 794;
  Best Local Similarity   27.2%;  Pred. No. 7.2;
  Matches   28;  Conservative   16;  Mismatches   47;  Indels   12;  Gaps    3;

Qy          2 QDLEQKMKVVENLQDDFDFNY-KTLKSQGDMQDLNGNN-------QSVTRQKMQQLEQML 53
              || | ::| ::  |: |   | ::|: |     |   |        ::  :||  || |
Db        157 QDTENELKKLQQTQEYFIIQYQESLRIQAQFAQLAQLNPQERLSRETALQQKQVSLEAWL 216

Qy         54 ----TALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRR 92
                  | |.|  :  :    |  :   |  : |:||  ||||
Db        217 QREAQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELIQWKRR 259


RESULT 9
G72593
hypothetical protein APE1216 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: G72593
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Taka
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: G72593
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-533 <KAW>
A;Cross-references: UNIPROT:Q9YCP2; DDBJ:AP000061; NID:g5104821; PIDN:BAA80205.1; PID:d
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1216

  Query Match             18.0%;  Score 86;  DB 2;  Length 533;
  Best Local Similarity   22.5%;  Pred. No. 6.9;
  Matches   23;  Conservative   27;  Mismatches   32;  Indels   20;  Gaps    3;

Qy          1 VQDLEQKMKVVEN----LQDDF--------------DFNYKTLKSQGDMQDLNGNNQSV 41
              |:||| ::  ||:     ::|                | : :  ::|  ::|||     |
Db        337 VEDLEARVGSVEDRLSQAEEDIDSLTTSLDSLRTELEDLSTRLAEAQASLEDLNTRLDQV 396

Qy         42 TRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTD 83
                :|||:| |   ::  :::  :|| | : :| :|::: :
Db        397 A-STLQQLQQRLATAEESLQALTEDLASLQAEVETLQQSIVE 437


RESULT 10
S33124
tpr protein - human
N;Alternate names: kinase-related transforming protein (tpr-met); protein with promoter
C;Species: Homo sapiens (man)
C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C;Accession: S33124; S23740; S00928; G01185
R;Mitchell, P.J.; Cooper, C.S.
Oncogene 7, 2329-2333, 1992
A;Title: The human tpr gene encodes a protein of 2094 amino acids that has extensive co
A;Reference number: S33124; MUID:93064711; PMID:1437155
A;Accession: S33124
```

```
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-2094 <MIT>
A;Cross-references: UNIPROT:Q15624; UNIPROT:Q9UE33; EMBL:X66397; NID:g633225
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991
R;Mitchell, P.J.; Cooper, C.S.
Oncogene 7, 383-388, 1992
A;Title: Nucleotide sequence analysis of human tpr cDNA clones.
A;Reference number: S23740; MUID:92195670; PMID:1549355
A;Accession: S23740
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-725,'L' <MI2>
A;Cross-references: EMBL:X63105; NID:g37257; PIDN:CAA44819.1; PID:g37258
R;King, H.W.S.; Tempest, P.R.; Merrifield, K.R.; Rance, A.J.
Oncogene 2, 617-619, 1988
A;Title: Tpr homologues activate met and raf.
A;Reference number: S00928; MUID:88262257; PMID:3387099
A;Accession: S00928
A;Molecule type: mRNA
A;Residues: 1-31,'R',33-142 <KIN>
A;Cross-references: EMBL:Y00672; NID:g37255; PIDN:CAA68681.1; PID:g37256
R;Greco, A.
submitted to the EMBL Data Library, December 1995
A;Reference number: H00592
A;Accession: G01185
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 144-228 <GRE>
A;Cross-references: EMBL:X94208; NID:g1296797; PIDN:CAA63904.1; PID:g1296798
C;Genetics:
A;Gene: GDB:TPR
A;Cross-references: GDB:128821; OMIM:189940
A;Map position: 1q25-1q25
A;Introns: 177/3

  Query Match             17.9%;  Score 85.5;  DB 2;  Length 2094;
  Best Local Similarity   26.5%;  Pred. No. 33;
  Matches   26;  Conservative   24;  Mismatches   21;  Indels   27;  Gaps    4;

Qy          3 DLEQKMKV---VENLQDDFDFNYKTLKSQ-----------GDMQDLNGNNQSVTRQKMQ 47
              |:::|:|    |: :   :   |: ||:|           || |:     | |: |:||
Db       1423 DIQEKVKTITQVKKIGRRYKTQYEELKAQQDKVMETSAQSSGDHQE-----QHVSVQEMQ 1477

Qy         48 QLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEE 85
              :|:: |   :   :|: |:       :| :||||:::|
Db       1478 ELKETLNQAETKSKSLESQ-------VENLQKTLSEKE 1508


RESULT 11
G75219
hypothetical protein PAB2166 - Pyrococcus abyssi (strain Orsay)
C;Species: Pyrococcus abyssi
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: G75219
R;anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A;Reference number: A75001
A;Accession: G75219
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-217 <KAW>
A;Cross-references: UNIPROT:Q9V1Z2; GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAB4920
A;Experimental source: strain Orsay
C;Genetics:
A;Gene: PAB2166

  Query Match             17.7%;  Score 84.5;  DB 2;  Length 217;
  Best Local Similarity   24.0%;  Pred. No. 3.6;
  Matches   23;  Conservative   31;  Mismatches   31;  Indels   11;  Gaps    3;

Qy          1 VQDLEQKMKVVENLQDDFD-FNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQM 59
              . :::|::|:: :| |:|||| : :| :| :: : |: :: || :  :| :
Db       132 LEELKRKVRRLEVLEDDFDHLKEQLIKQEGQLE--------MAREYIKLLEGVRHDIDNI 183

Qy         60 RRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              : :: : |:| |   : ::| : || |   :|  |
Db       184 KANVETLLSGYLE--DVIRKIVNDELNARGLKRTPI 217


RESULT 12
T24728
hypothetical protein T09A5.10 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T24728
R;Lightning, J.
submitted to the EMBL Data Library, August 1994
A;Reference number: Z19928
A;Accession: T24728
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-821 <WIL>
A;Cross-references: UNIPROT:P45970; EMBL:Z36753; PIDN:CAA85342.1; GSPDB:GN00020; CESP:T
A;Experimental source: clone T09A5
C;Genetics:
A;Gene: CESP:T09A5.10
A;Map position: 2
A;Introns: 47/2; 253/2; 604/3; 643/1; 712/3

  Query Match             17.7%;  Score 84.5;  DB 2;  Length 821;
  Best Local Similarity   23.9%;  Pred. No. 15;
  Matches   27;  Conservative   22;  Mismatches   31;  Indels   33;  Gaps    3;

Qy          2 QDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNG--------------NNQSVTRQKMQ 47
              ::|  |:| |       : | | :|   :::|||                |: :|
Db       318 ENLTAKLKTVT------EDNGKLMKQARELRDLNDYEFARFRQQEQELTETLRATQDQMA 371

Qy         48 QLEQMLTALDQMRRSIVSELAGLLSAMEYV-------------QKTLTDEELA 87
              |:: || :::: | |: ||    | :::| :              ||:  ||||
Db       372 DLQEQLTGVEKIRASLKSENESLSASVEELSVASLRNKQDADNSKTMLSEELA 424


RESULT 13
JC5837
364K Golgi complex-associated protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004
C;Accession: JC5837
R;Toki, C.; Fujiwara, T.; Sohda, M.; Hong, H.S.; Misumi, Y.; Ikehara, Y.
Cell Struct. Funct. 22, 565-577, 1997
A;Title: Identification and characterization of rat 364-kDa Golgi-associated protein re
A;Reference number: JC5837; MUID:98093490; PMID:9431462
A;Accession: JC5837
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-3187 <TOK>
A;Cross-references: UNIPROT:Q63714; DDBJ:D25543; NID:g516825; PIDN:BAA05026.1; PID:g516
C;Comment: This protein plays a role in the formation and maintenance of the characteri
C;Superfamily: giantin
F;49-549,624-1176,1238-1707,1763-3114/Domain: coiled-coil leucine zipper #status predic
F;3165-3187/Domain: membrane anchor #status predicted <MAD>

  Query Match             17.5%;  Score 83.5;  DB 2;  Length 3187;
  Best Local Similarity   25.5%;  Pred. No. 76;
  Matches   26;  Conservative   17;  Mismatches   40;  Indels   19;  Gaps    2;

Qy          3 DLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQ---- 58
              :|::| |:: |||    |       || |       :  : |||:|: |::: ||
Db       150 ELQEKEKLISNLQAQLDQAQSEQASQLDKSSAEMEDFILMRQKLQEKEELIGALQTQLSQ 209

Qy         59 ------------MRRSIVSE---LAGLLSAMEYVQKTLTDEE 85
                          |:| :     | |    :: :|| ||  |
Db       210 TQAEQAAQKLRVMQRKLEEHEEALLGRAQVVDLLQKELTSAE 251
```

```
RESULT 14
S73852
hypothetical protein MG218 homolog F10_orf1818 - Mycoplasma pneumoniae (strain ATCC 2934
C;Species: Mycoplasma pneumoniae
A;Variety: ATCC 29342
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S73852
R;Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A;Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae
A;Reference number: S73327; MUID:97105885; PMID:8948633
A;Accession: S73852
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1818 <HIM>
A;Cross-references: UNIPROT:P75471; EMBL:AE000051; GB:U00089; NID:g1674211; PIDN:AAB9617
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C;Genetics:
A;Genetic code: SGC3
C;Superfamily: Mycoplasma genitalium hypothetical protein MG218

  Query Match           17.2%;  Score 82;  DB 1;  Length 1818;
  Best Local Similarity  27.5%;  Pred. No. 56;
  Matches   28;  Conservative   21;  Mismatches   33;  Indels   20;  Gaps    4;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTL-KSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQM 59
              | :| |: | : : |   |   : | ::|  :|:|:    | |::| | ||  :   ||
Db       1291 VSELRQEKKQLLSAQKQVDDKSRLLEQNQRHLQNLS----SETKKKRQSLEHDINKFDQR 1346

Qy         60 RRSIVS--------------ELAGLLSAMEYVQKTLTDEELA 87
              |:  ||              || |:|  :  ::||  ::| :
Db       1347 RKEAVSSILNSHKKLKQKEGELQGILQKLS-LKKTQIEQEFS 1387


RESULT 15
B71203
hypothetical protein PH1895 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 09-Jul-2004
C;Accession: B71203
R;Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: B71203
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-217 <KAW>
A;Cross-references: UNIPROT:O59525; GB:AP000007; NID:g3236134; PIDN:BAA31017.1; PID:g325
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH1895

  Query Match           17.1%;  Score 81.5;  DB 2;  Length 217;
  Best Local Similarity  23.5%;  Pred. No. 6.4;
  Matches   20;  Conservative   30;  Mismatches   24;  Indels   11;  Gaps    3;

Qy          1 VQDLEQKMKVVENLQDDFD-FNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQM 59
              :::|::|:: :| ::|||    : :| :| ::        : |: :: || :   :| :
Db        132 LEELKKKIRRLEVIEDDFDHLKEQLIKQEGQLE--------MAREYIKLLEGVRHDIDNI 183

Qy         60 RRSIVSELAGLLSAMEYVQKTLTDE 84
              : :| : |:| :   : |:| : ||
Db        184 KANIETLLSGYIE--DVVRKIVNDE 206


Search completed: May 25, 2005, 17:45:23
Job time : 12.9623 secs
```

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:15:30 ; Search time 39.5283 Seconds
(without alignments)
1230.701 Million cell updates/sec

Title: US-10-090-185-21
Perfect score: 477
Sequence: 1 VQDLEQKMKVVENLQDDFDF.....YVQKTLTDEELADWKRRPEI 95

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	466	97.7	770	1	STA3_HUMAN	P40763 homo sapien
2	466	97.7	770	1	STA3_MOUSE	P42227 mus musculu
3	466	97.7	770	1	STA3_RAT	P52631 rattus norv
4	459	96.2	770	1	STA3_BOVIN	P61635 bos taurus
5	451	94.5	771	2	Q6DV79	Q6dv79 gallus gall
6	440.5	92.3	108	2	Q704W6	Q704w6 bos taurus
7	432	90.6	769	2	Q9PVX8	Q9pvx8 xenopus lae
8	427	89.5	766	2	Q7ZXK3	Q7zxk3 xenopus lae
9	381	79.9	765	2	Q6DVF3	Q6dvf3 oryzias lat
10	381	79.9	785	2	Q6GUE7	Q6gue7 oryzias lat
11	379	79.5	414	2	Q7ZTS5	Q7zts5 brachydanio
12	379	79.5	786	2	Q6NV46	Q6nv46 brachydanio
13	379	79.5	806	2	O93599	O93599 brachydanio
14	375	78.6	767	2	O13133	O13133 oncorhynchu
15	374	78.4	764	2	Q90Y16	Q90y16 tetraodon f
16	173	36.3	748	1	STA4_HUMAN	Q14765 homo sapien
17	170	35.6	712	2	Q99K94	Q99k94 mus musculu
18	170	35.6	749	1	STA1_MOUSE	P42225 mus musculu
19	170	35.6	749	2	Q8C3V4	Q8c3v4 mus musculu
20	170	35.6	749	2	Q8C497	Q8c497 mus musculu
21	170	35.6	749	2	Q9D323	Q9d323 mus musculu
22	170	35.6	755	2	Q8C8M3	Q8c8m3 mus musculu
23	169	35.4	757	2	Q764M5	Q764m5 sus scrofa
24	168	35.2	712	2	Q6P6Q7	Q6p6q7 rattus norv
25	168	35.2	737	2	Q90Y15	Q90y15 tetraodon f
26	168	35.2	749	2	Q9QXK0	Q9qxk0 rattus norv
27	168	35.2	1165	2	Q7TP57	Q7tp57 rattus norv
28	163	34.2	657	2	Q8AW24	Q8aw24 brachydanio
29	162.5	34.1	718	2	Q801Y2	Q801y2 carassius a
30	161	33.8	748	2	Q66HB2	Q66hb2 rattus norv
31	160.5	33.6	749	1	STA4_MOUSE	P42228 mus musculu
32	158	33.1	750	1	STA1_HUMAN	P42224 homo sapien
33	158	33.1	750	2	Q68D00	Q68d00 homo sapien
34	156	32.7	528	2	Q8JFU8	Q8jfu8 brachydanio
35	155	32.5	751	2	Q8JGN0	Q8jgn0 xenopus lae
36	154	32.3	651	2	Q7ZZ53	Q7zz53 brachydanio
37	154	32.3	667	2	Q8AW20	Q8aw20 brachydanio
38	153.5	32.2	652	2	Q7ZZ77	Q7zz77 brachydanio
39	153	32.1	553	2	Q8JFU7	Q8jfu7 brachydanio
40	147	30.8	754	2	O13131	O13131 oncorhynchu
41	146	30.6	749	2	O93598	O93598 brachydanio
42	146	30.6	749	2	Q6P943	Q6p943 brachydanio
43	134	28.1	754	2	O13132	O13132 oncorhynchu
44	123	25.8	758	2	Q90Y17	Q90y17 tetraodon f
45	121	25.4	1153	2	Q8JFS5	Q8jfs5 brachydanio

ALIGNMENTS

RESULT 1
STA3_HUMAN
ID STA3_HUMAN STANDARD; PRT; 770 AA.
AC P40763; O14916; Q9BW54;
DT 01-FEB-1995 (Rel. 31, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Signal transducer and activator of transcription 3 (Acute-phase
DE response factor).
GN Name=STAT3; Synonyms=APRF;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Placenta;
RX MEDLINE=94208062; PubMed=7512451; DOI=10.1016/0092-8674(94)90235-6;
RA Akira S., Nishio Y., Inoue M., Wang X.-J., Wei S., Matsusaka T.,
RA Yoshida K., Sudo T., Naruto M., Kishimoto T.;
RT "Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-
RT related transcription factor involved in the gp130-mediated signaling
RT pathway.";
RL Cell 77;63-71(1994).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=98296260; PubMed=9630560; DOI=10.1016/S0378-1119(98)00185-1;
RA Della Pietra L., Bressan A., Pezzotti A., Serlupi-Crescenzi O.;
RT "Highly conserved amino-acid sequence between murine STAT3 and a
RT revised human STAT3 sequence.";
RL Gene 213:119-124(1998).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM DEL-701), AND VARIANT ILE-143.
RA Rieder M.J., Daniels R.L., da Ponte S.H., Hastings N.C., Ahearn M.O.,
RA Rajkumar N., Yi Q., Nickerson D.A.;
RT "SeattleSNPs. NHLBI HL66682 program for genomic applications, UW-
RT FHCRC, Seattle, WA (URL: http://pga.gs.washington.edu).";
RL Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND DEL-701).
RC TISSUE=Kidney, and Pancreas;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,

```
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   SEQUENCE OF 564-704 FROM N.A.
RC   TISSUE=Liver;
RA   Della Pietra L., Bressan A., Pezzotti A.R., Serlupi-Crescenzi O.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   PHOSPHORYLATION ON SERINE.
RX   MEDLINE=95215843; PubMed=7701321;
RA   Zhang X., Blenis J., Li H.-C., Schindler C., Chen-Kiang S.;
RT   "Requirement of serine phosphorylation for formation of STAT-promoter
RT   complexes.";
RL   Science 267:1990-1994(1995).
RN   [7]
RP   INTERACTION WITH NCOA1.
RX   PubMed=11773079; DOI=10.1074/jbc.M111486200;
RA   Giraud S., Bienvenu F., Avril S., Gascan H., Heery D.M., Coqueret O.;
RT   "Functional interaction of STAT3 transcription factor with the
RT   coactivator NcoA/SRC1a.";
RL   J. Biol. Chem. 277:8004-8011(2002).
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes.
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P40763-1; Sequence=Displayed;
CC       Name=Del-701;
CC         IsoId=P40763-2; Sequence=VSP_010474;
CC   -!- TISSUE SPECIFICITY: Heart, brain, placenta, lung, liver, skeletal
CC       muscle, kidney and pancreas.
CC   -!- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,
CC       LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation
CC       is important for the formation of stable DNA-binding STAT3
CC       homodimers and maximal transcriptional activity.
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L29277; AAA58374.1; -.
DR   EMBL; AJ012463; CAA10032.1; -.
DR   EMBL; AY572796; AAS66986.1; -.
DR   EMBL; BC000627; AAH00627.1; -.
DR   EMBL; BC014482; AAH14482.1; -.
DR   EMBL; AF029311; AAB84254.1; -.
DR   PIR; A54444; A54444.
DR   HSSP; P42227; 1BG1.
DR   TRANSFAC; T01493; -.
DR   Genew; HGNC:11364; STAT3.
DR   H-InvDB; HIX0013840; -.
DR   MIM; 102582; -.
DR   GO; GO:0005737; C:cytoplasm; TAS.
DR   GO; GO:0005634; C:nucleus; TAS.
DR   GO; GO:0005062; F:hematopoietin/interferon-class (D200-domain. . .; TAS.
DR   GO; GO:0003700; F:transcription factor activity; TAS.
DR   GO; GO:0006928; P:cell motility; TAS.
DR   GO; GO:0007259; P:JAK-STAT cascade; TAS.
DR   GO; GO:0000122; P:negative regulation of transcription from P. . .; TAS.
DR   GO; GO:0007399; P:neurogenesis; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
KW   Activator; Alternative splicing; DNA-binding; Nuclear protein;
KW   Phosphorylation; Polymorphism; SH2 domain; Transcription regulation.
FT   DOMAIN       580    670       SH2.
FT   MOD_RES      705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES      727    727       Phosphoserine (By similarity).
FT   VARSPLIC     701    701       Missing (in isoform Del-701).
FT                                 /FTId=VSP_010474.
FT   VARIANT       32     32       Q -> K (in dbSNP:1803125).
FT                                 /FTId=VAR_018683.
FT   VARIANT      143    143       M -> I.
FT                                 /FTId=VAR_018679.
FT   CONFLICT     288    288       Q -> H (in Ref. 1).
FT   CONFLICT     460    460       P -> S (in Ref. 1).
FT   CONFLICT     548    548       K -> N (in Ref. 1).
FT   CONFLICT     561    561       F -> Y (in Ref. 1).
FT   CONFLICT     667    667       V -> L (in Ref. 1).
FT   CONFLICT     730    730       T -> A (in Ref. 1).
SQ   SEQUENCE   770 AA;  88067 MW;  6C00632211C8012D CRC64;
```

Query Match 97.7%; Score 466; DB 1; Length 770;
Best Local Similarity 97.9%; Pred. No. 2.4e-31;
Matches 93; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 214

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              ||||||||||||||||||||||||||||||||||| :|
Db        215 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249
```

```
RESULT 2
STA3_MOUSE
ID   STA3_MOUSE     STANDARD;      PRT;   770 AA.
AC   P42227;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Signal transducer and activator of transcription 3 (Acute-phase
DE   response factor).
GN   Name=Stat3; Synonyms=Aprf;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A), AND SEQUENCE OF 154-158; 181-185
RP   AND 632-640.
RC   STRAIN=BALB/c; TISSUE=Liver;
RX   MEDLINE=94208062; PubMed=7512451; DOI=10.1016/0092-8674(94)90235-6;
RA   Akira S., Nishio Y., Inoue M., Wang X.J., Shi W., Matsusaka T.,
RA   Yoshida K., Sudo T., Naruto M., Kishimoto T.;
RT   "Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-
RT   related transcription factor involved in the gp130-mediated signaling
RT   pathway.";
RL   Cell 77:63-71(1994).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
```

```
RC   TISSUE=Thymus;
RX   MEDLINE=94188718; PubMed=8140422;
RA   Zhong Z., Wen Z., Darnell J.E. Jr.;
RT   "Stat3: a STAT family member activated by tyrosine phosphorylation in
RT   response to epidermal growth factor and interleukin-6.";
RL   Science 264:95-98(1994).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   TISSUE=Brain;
RX   MEDLINE=95014185; PubMed=7523373;
RA   Raz R., Durbin J.E., Levy D.E.;
RT   "Acute phase response factor and additional members of the interferon-
RT   stimulated gene factor 3 family integrate diverse signals from
RT   cytokines, interferons, and growth factors.";
RL   J. Biol. Chem. 269:24391-24395(1994).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3B).
RC   STRAIN=BALB/c, and C57BL/6; TISSUE=Liver;
RX   MEDLINE=96016116; PubMed=7568080;
RA   Schaefer T.S., Sanders L.K., Nathans D.;
RT   "Cooperative transcriptional activity of Jun and Stat3 beta, a short
RT   form of Stat3.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:9097-9101(1995).
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   STRAIN=129/SvJ;
RX   PubMed=11161808; DOI=10.1006/geno.2000.6433;
RA   Miyoshi K., Cui Y., Riedlinger G., Lehoczky J., Zon L., Oka T.,
RA   Dewar K., Hennighausen L.;
RT   "Structure of the mouse stat 3/5 locus: evolution from Drosophila to
RT   zebrafish to mouse.";
RL   Genomics 71:150-155(2001).
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   STRAIN=C57BL/6J, and NOD/LtJ;
RA   Davoodi-Semiromi A., She J.-X.;
RT   "A mutant Stat5b with weaker DNA binding defines a key defective
RT   pathway in non-obese diabetic (NOD) mice.";
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   STRAIN=FVB/N; TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [8]
RP   PHOSPHORYLATION SITE SER-727, AND MUTAGENESIS.
RX   MEDLINE=95354205; PubMed=7543024; DOI=10.1016/0092-8674(95)90311-9;
RA   Wen Z., Zhong Z., Darnell J.E. Jr.;
RT   "Maximal activation of transcription by Stat1 and Stat3 requires both
RT   tyrosine and serine phosphorylation.";
RL   Cell 82:241-250(1995).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 136-716.
RX   MEDLINE=98334373; PubMed=9671298; DOI=10.1038/28101;
RA   Becker S., Groner B., Mueller C.W.;
RT   "Three-dimensional structure of the Stat3beta homodimer bound to
RT   DNA.";
RL   Nature 394:145-151(1998).
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes. STAT3B interacts with the N-terminal
CC       part of JUN to activate such promoters in a cooperative way.
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=Stat3A;
CC         IsoId=P42227-1; Sequence=Displayed;
CC       Name=Stat3B;
CC         IsoId=P42227-2; Sequence=VSP_006287;
CC       Name=Del-701;
CC         IsoId=P42227-3; Sequence=VSP_010475;
CC   -!- TISSUE SPECIFICITY: STAT3A is seen in the liver, spleen, and
CC       kidney. STAT3B is also detected in the liver, although in a much
CC       less abundant manner.
CC   -!- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,
CC       LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation
CC       is important for the formation of stable DNA-binding STAT3
CC       homodimers and maximal transcriptional activity (By similarity).
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L29278; AAA37254.1; -.
DR   EMBL; U06922; AAA19452.1; -.
DR   EMBL; U08378; AAA56668.1; -.
DR   EMBL; U30709; AAC52612.1; -.
DR   EMBL; AF246978; AAL59017.1; -.
DR   EMBL; AY299489; AAQ75418.1; -.
DR   EMBL; AY299490; AAQ75419.1; -.
DR   EMBL; BC003806; AAH03806.1; -.
DR   PIR; I49508; I49508.
DR   PDB; 1BG1; X-ray; A=1-722.
DR   TRANSFAC; T01574; -.
DR   MGD; MGI:103038; Stat3.
DR   GO; GO:0005737; C:cytoplasm; IDA.
DR   GO; GO:0005634; C:nucleus; IDA.
DR   GO; GO:0005886; C:plasma membrane; IDA.
DR   GO; GO:0003677; F:DNA binding; IDA.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   GO; GO:0016563; F:transcriptional activator activity; IDA.
DR   GO; GO:0007259; P:JAK-STAT cascade; IDA.
DR   GO; GO:0006357; P:regulation of transcription from Pol II pro. . .; IDA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
KW   3D-structure; Activator; Acute phase; Alternative splicing;
KW   Direct protein sequencing; DNA-binding; Nuclear protein;
KW   Phosphorylation; SH2 domain; Transcription regulation.
FT   DOMAIN      580    670       SH2.
FT   MOD_RES     705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES     727    727       Phosphoserine.
```

```
FT   VARSPLIC    716    770       TTCSNTIDLPMSPRTLDSLMQFGNNGEGAEPSAGGQFESLT
FT                                FDMDLTSECATSPM -> FIDAVWK (in isoform
FT                                Stat3B).
FT                                /FTId=VSP_006287.
FT   VARSPLIC    701    701       Missing (In isoform Del-701).
FT                                /FTId=VSP_010475.
FT   MUTAGEN     727    727       S->A: Decreased transcriptional
FT                                activation.
FT   CONFLICT     16     16       E -> K (in Ref. 2).
FT   CONFLICT     25     25       S -> T (in Ref. 2 and 4).
FT   CONFLICT    394    394       M -> I (in Ref. 1).
FT   HELIX       139    180
FT   TURN        181    182
FT   TURN        197    198
FT   HELIX       199    237
FT   TURN        238    238
FT   HELIX       239    251
FT   TURN        252    253
FT   HELIX       261    290
FT   TURN        294    295
FT   TURN        297    301
FT   HELIX       302    320
FT   STRAND      321    328
FT   TURN        330    331
FT   TURN        333    334
FT   TURN        336    337
FT   STRAND      338    340
FT   TURN        341    342
FT   STRAND      345    351

  Query Match             97.7%;  Score 466;  DB 1;  Length 770;
  Best Local Similarity   97.9%;  Pred. No. 2.4e-31;
  Matches   93;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 214

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              ||||||||||||||||||||||||||||||||| :|
Db        215 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249


RESULT 3
STA3_RAT
ID   STA3_RAT       STANDARD;      PRT;   770 AA.
AC   P52631;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Signal transducer and activator of transcription 3.
GN   Name=Stat3;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=96102059; PubMed=8530402; DOI=10.1074/jbc.270.50.29998;
RA   Ripperger J.A., Fritz S., Richter K., Hocke G.M., Lottspeich F.,
RA   Fey G.H.;
RT   "Transcription factors Stat3 and Stat5b are present in rat liver
RT   nuclei late in an acute phase response and bind interleukin-6 response
RT   elements.";
RL   J. Biol. Chem. 270:29998-30006(1995).
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes.
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation (By similarity).
CC   -!- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,
CC       LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation
CC       is important for the formation of stable DNA-binding STAT3
CC       homodimers and maximal transcriptional activity (By similarity).
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X91810; CAA62920.1; -.
DR   HSSP; P42227; 1BG1.
DR   RGD; 3772; Stat3.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
KW   Activator; DNA-binding; Nuclear protein; Phosphorylation; SH2 domain;
KW   Transcription regulation.
FT   DOMAIN      580    670       SH2.
FT   MOD_RES     705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES     727    727       Phosphoserine (By similarity).
SQ   SEQUENCE   770 AA;  88039 MW;  D74A0C76954754ED CRC64;

  Query Match             97.7%;  Score 466;  DB 1;  Length 770;
  Best Local Similarity   97.9%;  Pred. No. 2.4e-31;
  Matches   93;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 214

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              ||||||||||||||||||||||||||||||||| :|
Db        215 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249


RESULT 4
STA3_BOVIN
ID   STA3_BOVIN     STANDARD;      PRT;   770 AA.
AC   P61635;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Signal transducer and activator of transcription 3.
GN   Name=STAT3;
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Mammary gland;
RA   Seyfert H.M., Wheeler T.T., Moolenaar A., Pitra C.;
RT   "The STAT5B-encoding gene was flipped across the STAT3/STAT5A-locus
RT   during ruminant evolution.";
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes (By similarity).
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
```

```
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation (By similarity).
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ620655; CAF06182.1; -.
DR   PROSITE; PS50001; SH2; 1.
KW   Activator; DNA-binding; Nuclear protein; Phosphorylation; SH2 domain;
KW   Transcription regulation.
FT   DOMAIN      580    670       SH2.
FT   MOD_RES     705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES     727    727       Phosphoserine (By similarity).
SQ   SEQUENCE   770 AA;  87974 MW;  9CEB147C73E83274 CRC64;

  Query Match             96.2%;  Score 459;  DB 1;  Length 770;
  Best Local Similarity   96.8%;  Pred. No. 9.4e-31;
  Matches   92;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        155 VQDLEQKMKVVENLLDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 214

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||||| :|
Db        215 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249


RESULT 5
Q6DV79
ID   Q6DV79      PRELIMINARY;      PRT;   771 AA.
AC   Q6DV79;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Signal transducer and activator of transcription 3.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zhou G.Y., Leung F.C.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY641397; AAT64887.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   771 AA;  88174 MW;  71AC855C5DEC03E2 CRC64;

  Query Match             94.5%;  Score 451;  DB 2;  Length 771;
  Best Local Similarity   94.7%;  Pred. No. 4.5e-30;
  Matches   90;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 214

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              | |||||||||||||||| | |||||||||| :|
Db        215 RGIVSELAGLLSAMEYVQKMLADEELADWKRRQQI 249


RESULT 6
Q704W6
ID   Q704W6      PRELIMINARY;      PRT;   108 AA.
AC   Q704W6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Signal transducer and activator of transcription 3 (Fragment).
GN   Name=stat3;
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Mammary gland;
RA   Seyfert H.M., Wheeler T.T., Moolenaar A., Pitra C.;
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ620660; CAF06187.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF01017; STAT_alpha; 1.
FT   NON_TER       1      1
FT   NON_TER     108    108
SQ   SEQUENCE   108 AA;  12552 MW;  9961C28E59085803 CRC64;

  Query Match             92.3%;  Score 440.5;  DB 2;  Length 108;
  Best Local Similarity   96.8%;  Pred. No. 4.1e-30;
  Matches   90;  Conservative    1;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          3 DLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRS 62
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db          1 DLEQKMKVVENLQDDFDFNYKTLKSQ-DMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRS 59

Qy         63 IVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |||||||||||||||||||||||||||||| :|
Db         60 IVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 92


RESULT 7
Q9PVX8
ID   Q9PVX8      PRELIMINARY;      PRT;   769 AA.
AC   Q9PVX8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat 3.
GN   Name=stat 3;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20107399; PubMed=10642787; DOI=10.1006/dbio.1999.9518;
RA   Nishinakamura R., Matsumoto Y., Matsuda T., Ariizumi T., Heike T.,
RA   Asashima M., Yokota T.;
```

```
RT   "Activation of Stat3 by cytokine receptor gp130 ventralizes Xenopus
RT   embryos independent of BMP-4.";
RL   Dev. Biol. 216:481-490(1999).
DR   EMBL; AB017701; BAA86061.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   769 AA;  87974 MW;  0905C03263303069 CRC64;

  Query Match             90.6%;  Score 432;  DB 2;  Length 769;
  Best Local Similarity   88.4%;  Pred. No. 1.9e-28;
  Matches   84;  Conservative    7;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              |||||||||||||||||||||||||||| |: :|||||||||||||||||||||||||:|
Db        155 VQDLEQKMKVVENLQDDFDFNYKTLKSQSDLSELNGNNQSVTRQKMQQLEQMLTALDQLR 214

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |:|:|:|| |||||||||||||||||||||| :|
Db        215 RTIISDLASLLSAMEYVQKTLTDEELADWKRRQQI 249


RESULT 8
Q7ZXK3
ID   Q7ZXK3      PRELIMINARY;      PRT;   766 AA.
AC   Q7ZXK3;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat3-A protein.
GN   Name=stat3-A;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC044717; AAH44717.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   766 AA;  87599 MW;  31018A3321CCEB9C CRC64;

  Query Match             89.5%;  Score 427;  DB 2;  Length 766;
  Best Local Similarity   87.4%;  Pred. No. 5e-28;
  Matches   83;  Conservative    7;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMR 60
              |||||||||||||||||||||||||||| |: :|||||||||||||||||||||||||:|
Db        155 VQDLEQKMKVVENLQDDFDFNYKTLKSQSDLSELNGNNQSVTRQKMQQLEQMLTALDQLR 214

Qy         61 RSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              |:|:|:|| |||||||||| |||||||||||| :|
Db        215 RTIISDLASLLSAMEYVQKNLTDEELADWKRRQQI 249


RESULT 9
Q6DVF3
ID   Q6DVF3      PRELIMINARY;      PRT;   765 AA.
AC   Q6DVF3;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Signal transducer and activation of transcription factor 3.
OS   Oryzias latipes (Medaka fish) (Japanese ricefish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Liu R., Hong Y.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY639947; AAT64912.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
```

```
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   765 AA;  87566 MW;  F5D01408748EC703 CRC64;

  Query Match             79.9%;  Score 381;  DB 2;  Length 765;
  Best Local Similarity   77.3%;  Pred. No. 4.1e-24;
  Matches   75;  Conservative   14;  Mismatches    6;  Indels    2;  Gaps    2;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLNGNNQ-SVTRQKMQQLEQMLTALDQ 58
              |||:|||||::||||||||||||||||:: |||||:| : ||||| ||||||:||||
Db        154 VQDMEQKMKMLENLQDDFDFNYKTLKSQGELNQDLNGNSQAAATRQKMAQLEQMLSALDQ 213

Qy         59 MRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              :|| ||:|: |||:||:|||| |||||||||||| :|
Db        214 LRRQIVTEMGGLLTAMDYVQKNLTDEELADWKRRQQI 250


RESULT 10
Q6GUE7
ID   Q6GUE7      PRELIMINARY;      PRT;   785 AA.
AC   Q6GUE7;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Signal transducer and activator of transcription 3 isoform 1.
OS   Oryzias latipes (Medaka fish) (Japanese ricefish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Liu R., Hong Y.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY641434; AAT46364.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   785 AA;  89643 MW;  81F231BDE27DE938 CRC64;

  Query Match             79.9%;  Score 381;  DB 2;  Length 785;
  Best Local Similarity   77.3%;  Pred. No. 4.2e-24;
  Matches   75;  Conservative   14;  Mismatches    6;  Indels    2;  Gaps    2;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLNGNNQ-SVTRQKMQQLEQMLTALDQ 58
              |||:|||||::||||||||||||||||:: |||||:| : ||||| ||||||:||||
Db        154 VQDMEQKMKMLENLQDDFDFNYKTLKSQGELNQDLNGNSQAAATRQKMAQLEQMLSALDQ 213

Qy         59 MRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              :|| ||:|: |||:||:|||| |||||||||||| :|
Db        214 LRRQIVTEMGGLLTAMDYVQKNLTDEELADWKRRQQI 250


RESULT 11
Q7ZTS5
ID   Q7ZTS5      PRELIMINARY;      PRT;   414 AA.
AC   Q7ZTS5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat3 protein.
GN   Name=stat3;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB; TISSUE=Whole body;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB; TISSUE=Whole body;
RA   Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC045276; AAH45276.1; -.
DR   HSSP; P42227; 1BG1.
DR   ZFIN; ZDB-GENE-980526-68; stat3.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
SQ   SEQUENCE   414 AA;  48253 MW;  0FFD1B509B7526BD CRC64;

  Query Match             79.5%;  Score 379;  DB 2;  Length 414;
  Best Local Similarity   77.3%;  Pred. No. 3.1e-24;
  Matches   75;  Conservative   14;  Mismatches    6;  Indels    2;  Gaps    2;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLNGNNQ-SVTRQKMQQLEQMLTALDQ 58
              |||:|||||::||||||||||||||||:: |||||:| : ||||| ||||||:||||
Db        154 VQDMEQKMKMLENLQDDFDFNYKTLKSAGELSQDLNGNSQAAATRQKMSQLEQMLSALDQ 213

Qy         59 MRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              :|| ||:|:|||||||::||| |||||||||||| :|
Db        214 LRRQIVTEMAGLLSAMDFVQKNLTDEELADWKRRQQI 250


RESULT 12
Q6NV46
ID   Q6NV46      PRELIMINARY;      PRT;   786 AA.
AC   Q6NV46;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Stat3 protein.
GN   Name=stat3;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC068320; AAH68320.1; -.
DR   ZFIN; ZDB-GENE-980526-68; stat3.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   786 AA;  90039 MW;  FC7371D0B0E5447E CRC64;

  Query Match             79.5%;  Score 379;  DB 2;  Length 786;
  Best Local Similarity   77.3%;  Pred. No. 6.2e-24;
  Matches   75;  Conservative   14;  Mismatches    6;  Indels    2;  Gaps    2;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLNGNNQ-SVTRQKMQQLEQMLTALDQ 58
              |||:|||||::|||||||||||||||| |:: ||||||:| : ||||| ||||||:||||
Db        154 VQDMEQKMKMLENLQDDFDFNYKTLKSAGELSQDLNGNSQAAATRQKMSQLEQMLSALDQ 213

Qy         59 MRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              :|| ||:|:|||||||::||| |||||||||||| :|
Db        214 LRRQIVTEMAGLLSAMDFVQKNLTDEELADWKRRQQI 250


RESULT 13
O93599
ID   O93599      PRELIMINARY;      PRT;   806 AA.
AC   O93599;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Transcription factor.
GN   Name=stat3;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Oates A.C.;
RL   Thesis (1998), University of Melbourne, Australia.
DR   EMBL; AJ005693; CAA06677.1; -.
DR   HSSP; P42227; 1BG1.
DR   ZFIN; ZDB-GENE-980526-68; stat3.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   806 AA;  92151 MW;  74BC4EA401C3C942 CRC64;

  Query Match             79.5%;  Score 379;  DB 2;  Length 806;
  Best Local Similarity   77.3%;  Pred. No. 6.4e-24;
  Matches   75;  Conservative   14;  Mismatches    6;  Indels    2;  Gaps    2;

Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLNGNNQ-SVTRQKMQQLEQMLTALDQ 58
              |||:|||||::|||||||||||||||| |:: ||||||:| : ||||| ||||||:||||
Db        154 VQDMEQKMKMLENLQDDFDFNYKTLKSAGELSQDLNGNSQAAATRQKMSQLEQMLSALDQ 213

Qy         59 MRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              :|| ||:|:|||||||::||| |||||||||||| :|
Db        214 LRRQIVTEMAGLLSAMDFVQKNLTDEELADWKRRQQI 250


RESULT 14
O13133
ID   O13133      PRELIMINARY;      PRT;   767 AA.
AC   O13133;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat3.
GN   Name=rbtStat3;
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8022;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Johnson M.C., Mourich D.V., Leong J.C.;
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U60333; AAB60926.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
SQ   SEQUENCE   767 AA;  87816 MW;  FED97740C74C3798 CRC64;
```

Query Match 78.6%; Score 375; DB 2; Length 767;
Best Local Similarity 75.3%; Pred. No. 1.3e-23;
Matches 73; Conservative 17; Mismatches 5; Indels 2; Gaps 2;

```
Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLNGNNQ-SVTRQKMQQLEQMLTALDQ 58
              |||:|||||::|||||||||:||||||:: ||:|||:| : ||||| |||||||:||||
Db        154 VQDMEQKMKMLENLQDDFDFHYKTLKSQGELSQDMNGNSQAAATRQKMSQLEQMLSALDQ 213

Qy         59 MRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              :|| ||:|:|||||||::||| |||:||||||||| :|
Db        214 LRRQIVTEMAGLLSAMDFVQKNLTDDELADWKRRQQI 250
```

RESULT 15
Q90Y16
ID Q90Y16 PRELIMINARY; PRT; 764 AA.
AC Q90Y16;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE STAT3.
OS Tetraodon fluviatilis (Puffer fish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetradontoidea; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=47145;
RN [1]
RP SEQUENCE FROM N.A.
RA Sung S.C., Fan T.J., Leu J.H., Chou C.M., Huang C.J.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF307106; AAL09415.1; -.
DR HSSP; P42227; 1BG1.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF01017; STAT_alpha; 1.
DR Pfam; PF02864; STAT_bind; 1.
DR Pfam; PF02865; STAT_int; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS50001; SH2; 1.
SQ SEQUENCE 764 AA; 87411 MW; E661FFE18BEFD8BE CRC64;

Query Match 78.4%; Score 374; DB 2; Length 764;
Best Local Similarity 76.3%; Pred. No. 1.6e-23;
Matches 74; Conservative 14; Mismatches 7; Indels 2; Gaps 2;

```
Qy          1 VQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLNGNNQ-SVTRQKMQQLEQMLTALDQ 58
              |||:|||||::|||||||||||||||||:: ||||||:| |  |||| |||||||:||||
Db        154 VQDMEQKMKMLENLQDDFDFNYKTLKSQGELSQDLNGNSQASAARQKMAQLEQMLSALDQ 213

Qy         59 MRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 95
              :|| ||:|: |||:||:|||| ||:||||||||| :|
Db        214 LRRQIVTEMGGLLTAMDYVQKNLTNEELADWKRRQQI 250
```

Search completed: May 25, 2005, 17:43:41
Job time : 44.5283 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:14:45 ; Search time 68.6634 Seconds
(without alignments)
726.619 Million cell updates/sec

Title: US-10-090-185-24
Perfect score: 669
Sequence: 1 IACIGGPPNICLDRLENWIT.....LNYQLKIKVCIDKDSGDVAA 129

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	669	100.0	129	4	AAY72856	Aay72856 Mouse Sta
2	669	100.0	185	4	AAY72855	Aay72855 Mouse Sta
3	669	100.0	223	4	AAY72854	Aay72854 Mouse Sta
4	669	100.0	271	4	AAY72841	Aay72841 Mouse Sta
5	669	100.0	720	5	AAE22055	Aae22055 Human Sta
6	669	100.0	769	5	ABB57164	Abb57164 Mouse isc
7	669	100.0	769	5	AAE22054	Aae22054 Human Sta
8	669	100.0	769	5	AAE22056	Aae22056 Human pro
9	669	100.0	770	2	AAR72082	Aar72082 Mouse Sta
10	669	100.0	770	2	AAR82995	Aar82995 Mouse liv
11	669	100.0	770	2	AAW03176	Aaw03176 Mouse STA
12	669	100.0	770	2	AAY03768	Aay03768 Human STA
13	669	100.0	770	3	AAB12377	Aab12377 N-termina
14	669	100.0	770	5	AAE14652	Aae14652 Murine ST
15	669	100.0	770	5	ABG69497	Abg69497 Human bai
16	669	100.0	770	6	ABU10476	Abu10476 Mouse STA
17	669	100.0	770	8	ADN04365	Adn04365 Antipsori
18	669	100.0	770	8	ADP54789	Adp54789 Human PRO
19	669	100.0	793	3	AAB58442	Aab58442 Lung canc
20	666	99.6	770	7	ADD44738	Add44738 Rat Prote
21	664	99.3	770	2	AAR82993	Aar82993 Human pla
22	664	99.3	770	4	AAB19964	Aab19964 Human sig
23	664	99.3	770	5	AAE15174	Aae15174 Human Sta
24	664	99.3	770	7	ADD44740	Add44740 Human Pro
25	571	85.4	228	4	AAY72861	Aay72861 Mouse Sta
26	571	85.4	229	4	AAY72850	Aay72850 Mouse Sta
27	571	85.4	229	4	AAY72863	Aay72863 Mouse Sta
28	571	85.4	229	4	AAY72862	Aay72862 Mouse Sta
29	571	85.4	252	4	AAY72846	Aay72846 Mouse Sta
30	554	82.8	229	4	AAY72860	Aay72860 Mouse Sta
31	494	73.8	96	4	AAY72857	Aay72857 Mouse Sta
32	493	73.7	213	4	AAY72851	Aay72851 Mouse Sta
33	493	73.7	236	4	AAY72847	Aay72847 Mouse Sta
34	436	65.2	268	4	AAY72844	Aay72844 Mouse Sta
35	436	65.2	582	2	AAW62996	Aaw62996 Human tru
36	436	65.2	582	6	ABU04748	Abu04748 Human exp
37	436	65.2	582	8	ADH57036	Adh57036 Truncated
38	436	65.2	680	6	ABR59713	Abr59713 Human sig
39	436	65.2	712	2	AAR72079	Aar72079 Human Sta
40	436	65.2	712	2	AAW03170	Aaw03170 Human STA
41	436	65.2	712	2	AAW62995	Aaw62995 Human Sta
42	436	65.2	712	6	ABU04747	Abu04747 Human exp
43	436	65.2	712	6	ABU04735	Abu04735 Human exp
44	436	65.2	712	6	ABU04745	Abu04745 Human exp
45	436	65.2	712	6	ABU04743	Abu04743 Human exp

ALIGNMENTS

RESULT 1
AAY72856
ID AAY72856 standard; protein; 129 AA.
XX
AC AAY72856;
XX
DT 31-MAY-2001 (first entry)
XX
DE Mouse Stat3 protein fragment #14 (249-377 amino acids).
XX
KW Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW cellular transformation; dysproliferative disease; cancer; psoriasis;
KW therapy.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT Region 94. .110
FT /note= "Stat3-c-Jun interaction region 2; corresponds to
FT 342-358 position of Stat3 protein"
XX
PN WO200116605-A2.
XX
PD 08-MAR-2001.
XX
PF 30-AUG-2000; 2000WO-US023822.
XX
PR 31-AUG-1999; 99US-00387418.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI Zhang X, Horvath C, Wrzeszcynska MH, Darnell JE;
XX
DR WPI; 2001-226705/23.
XX
PT Identifying an agent for use in modulating the interaction between
PT transcription factor c-Jun and a Stat3 protein.
XX
PS Claim 65; Page 80-81; 86pp; English.
XX
CC The present sequence is mouse Stat3 protein fragment containing 249-377
CC amino acids of Stat3 protein. The invention relates to methods for
CC identifying interacting regions of transcription factors and methods for
CC identifying agents which modulates the interaction between a
CC transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC Stat-3, useful for modulating gene transcription e.g., cellular
CC transformation. These identifying agents are used in the treatment of
CC dysproliferative diseases and also for treating cancer and psoriasis. A

```
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 129 AA;

  Query Match             100.0%;  Score 669;  DB 4;  Length 129;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-69;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120

Qy        121 DKDSGDVAA 129
              |||||||||
Db        121 DKDSGDVAA 129


RESULT 2
AAY72855
ID   AAY72855 standard; protein; 185 AA.
XX
AC   AAY72855;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 protein fragment #13 (193-377 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          150. .166
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 80; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 193-377
CC   amino acids of Stat3 protein. The invention relates to methods for
CC   identifying interacting regions of transcription factors and methods for
CC   identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 185 AA;

  Query Match             100.0%;  Score 669;  DB 4;  Length 185;
  Best Local Similarity   100.0%;  Pred. No. 6.8e-69;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         57 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 116

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        117 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 176

Qy        121 DKDSGDVAA 129
              |||||||||
Db        177 DKDSGDVAA 185


RESULT 3
AAY72854
ID   AAY72854 standard; protein; 223 AA.
XX
AC   AAY72854;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 protein fragment #12 (155-377 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          188. .204
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 79; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 155-377
CC   amino acids of Stat3 protein. The invention relates to methods for
CC   identifying interacting regions of transcription factors and methods for
CC   identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 223 AA;

  Query Match             100.0%;  Score 669;  DB 4;  Length 223;
```

```
  Best Local Similarity   100.0%;  Pred. No. 8.7e-69;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         95 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 154

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 214

Qy        121 DKDSGDVAA 129
              |||||||||
Db        215 DKDSGDVAA 223


RESULT 4
AAY72841
ID   AAY72841 standard; protein; 271 AA.
XX
AC   AAY72841;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 protein fragment #2 (107-377 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          24. .48
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
FT   Region          236. .252
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 67-68; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 107-377
CC   amino acids of Stat3 protein. This Stat3 fragment showed strong binding
CC   to c-Jun protein in the cell extract. The invention relates to methods
CC   for identifying interacting regions of transcription factors and methods
CC   for identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 271 AA;

  Query Match           100.0%;  Score 669;  DB 4;  Length 271;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-68;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 202

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 262

Qy        121 DKDSGDVAA 129
              |||||||||
Db        263 DKDSGDVAA 271


RESULT 5
AAE22055
ID   AAE22055 standard; protein; 720 AA.
XX
AC   AAE22055;
XX
DT   25-JUL-2002  (first entry)
XX
DE   Human Stat3beta protein.
XX
KW   Human; signal transducer and activator of transcription 3; ischaemia;
KW   immune response; Stat3; coronary atherosclerosis; vascular occlusion;
KW   hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;
KW   inflammation; chronic obstructive pulmonary disease; cardiac arrest;
KW   insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;
KW   shock; chronic active hepatitis; adult respiratory distress syndrome;
KW   nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;
KW   Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;
KW   polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;
KW   proliferative disease; Grave's disease; ulcerative colitis; sarcoma;
KW   carcinoma; degenerative disorder; gene therapy; growth deficiency;
KW   cirrhosis; hypoproliferative disorder; lesion; Statbeta.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 713. .714
FT                   /note= "Encoded by ACA CCA TTC"
XX
PN   WO200220032-A1.
XX
PD   14-MAR-2002.
XX
PF   10-SEP-2001; 2001WO-US028254.
XX
PR   08-SEP-2000; 2000US-0231212P.
XX
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (UYSF-) UNIV SOUTH FLORIDA.
XX
PI   Yu H,  Pardoll D,  Jove R,  Dalton W;
XX
DR   WPI; 2002-362218/39.
DR   N-PSDB; AAD35066.
XX
PT   Modulating angiogenesis and an immune response in an individual, for
PT   treating a hypoxic or ischemic condition, comprises administering a
PT   compound that modulates the activity of a signal transducer and activator
PT   of transcription 3.
XX
PS   Disclosure; Page 87-89; 94pp; English.
XX
CC   The invention relates to a method of modulating angiogenesis and immune
CC   response. Method involves administering to an individual a compound that
CC   modulate the activity of signal transducer and activator of transcription
CC   3 (Stat3). Modulating angiogenesis is useful for treating or preventing
```

CC hypoxic or ischaemic condition or disorder which is the result of stroke,
CC ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,
CC tissue ischaemia in the lower extremities, infarction, trauma, vascular
CC occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,
CC chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,
CC epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest,
CC nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy
CC with neovascularisation. Suppressing an immune response is useful for
CC ameliorating a symptom of an autoimmune disease such as systemic lupus
CC erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,
CC Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,
CC mixed connective tissue disease, primary biliary cirrhosis, pernicious
CC anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,
CC gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,
CC idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's
CC disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune
CC infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and
CC dense deposit disease. The method is useful in preventing or treating
CC specific proliferative and oncogenic disease which includes sarcomas and
CC carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,
CC fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,
CC hypoproliferative disorders, physical trauma, lesions and wounds. The
CC method is also used in gene therapy. The present sequence is human
CC Stat3beta protein
XX
SQ Sequence 720 AA;

Query Match 100.0%; Score 669; DB 5; Length 720;
Best Local Similarity 100.0%; Pred. No. 4.2e-68;
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377
```

RESULT 6
ABB57164
ID ABB57164 standard; protein; 769 AA.
XX
AC ABB57164;
XX
DT 07-MAR-2002 (first entry)
XX
DE Mouse ischaemic condition related protein sequence SEQ ID NO:398.
XX
KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease.
XX
OS Mus musculus.
XX
PN WO200188188-A2.
XX
PD 22-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-JP004192.
XX
PR 18-MAY-2000; 2000JP-00145977.
XX
PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
DR WPI; 2002-034733/04.
DR N-PSDB; ABI99454.
XX
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or by
PT determining the expression profile of a gene group comprising these
PT genes.
XX
PS Claim 2; Page 1084-1087; 2690pp; English.
XX
CC The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (ABI99202 to ABI99912, encoding the
CC protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The expression
CC levels or expression profiles produced by these genes are used as an
CC indicator when screening for ischaemic condition-improving drugs or
CC therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR
CC primers for a mouse ischaemic condition related sequence, which are used
CC in the exemplification of the present invention
XX
SQ Sequence 769 AA;

Query Match 100.0%; Score 669; DB 5; Length 769;
Best Local Similarity 100.0%; Pred. No. 4.6e-68;
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377
```

RESULT 7
AAE22054
ID AAE22054 standard; protein; 769 AA.
XX
AC AAE22054;
XX
DT 25-JUL-2002 (first entry)
XX
DE Human Stat3 protein.
XX
KW Human; signal transducer and activator of transcription 3; ischaemia;
KW immune response; Stat3; coronary atherosclerosis; vascular occlusion;
KW hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;
KW inflammation; chronic obstructive pulmonary disease; cardiac arrest;
KW insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;
KW shock; chronic active hepatitis; adult respiratory distress syndrome;
KW nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;
KW Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;
KW polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;
KW proliferative disease; Grave's disease; ulcerative colitis; sarcoma;
KW carcinoma; degenerative disorder; gene therapy; growth deficiency;
KW cirrhosis; hypoproliferative disorder; lesion.
XX
OS Homo sapiens.
XX
PN WO200220032-A1.
XX
PD 14-MAR-2002.
XX
PF 10-SEP-2001; 2001WO-US028254.
XX

PR 08-SEP-2000; 2000US-0231212P.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
PA (UYSF-) UNIV SOUTH FLORIDA.
XX
PI Yu H, Pardoll D, Jove R, Dalton W;
XX
DR WPI; 2002-362218/39.
DR N-PSDB; AAD35065.
XX
PT Modulating angiogenesis and an immune response in an individual, for
PT treating a hypoxic or ischemic condition, comprises administering a
PT compound that modulates the activity of a signal transducer and activator
PT of transcription 3.
XX
PS Disclosure; Page 83-85; 94pp; English.
XX
CC The invention relates to a method of modulating angiogenesis and immune
CC response. Method involves administering to an individual a compound that
CC modulate the activity of signal transducer and activator of transcription
CC 3 (Stat3). Modulating angiogenesis is useful for treating or preventing
CC hypoxic or ischaemic condition or disorder which is the result of stroke,
CC ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,
CC tissue ischaemia in the lower extremities, infarction, trauma, vascular
CC occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,
CC chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,
CC epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest,
CC nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy
CC with neovascularisation. Suppressing an immune response is useful for
CC ameliorating a symptom of an autoimmune disease such as systemic lupus
CC erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,
CC Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,
CC mixed connective tissue disease, primary biliary cirrhosis, pernicious
CC anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,
CC gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,
CC idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's
CC disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune
CC infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and
CC dense deposit disease. The method is useful in preventing or treating
CC specific proliferative and oncogenic disease which includes sarcomas and
CC carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,
CC fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,
CC hypoproliferative disorders, physical trauma, lesions and wounds. The
CC method is also used in gene therapy. The present sequence is human Stat3
CC protein
XX
SQ Sequence 769 AA;

Query Match 100.0%; Score 669; DB 5; Length 769;
Best Local Similarity 100.0%; Pred. No. 4.6e-68;
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy 61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy 121 DKDSGDVAA 129
|||||||||
Db 369 DKDSGDVAA 377

RESULT 8
AAE22056
ID AAE22056 standard; protein; 769 AA.
XX
AC AAE22056;
XX
DT 25-JUL-2002 (first entry)
XX
DE Human protein related to angiogenesis regulation.
XX
KW Human; signal transducer and activator of transcription 3; ischaemia;
KW immune response; Stat3; coronary atherosclerosis; vascular occlusion;
KW hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;
KW inflammation; chronic obstructive pulmonary disease; cardiac arrest;
KW insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;
KW shock; chronic active hepatitis; adult respiratory distress syndrome;
KW nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;
KW Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;
KW polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;
KW proliferative disease; Grave's disease; ulcerative colitis; sarcoma;
KW carcinoma; degenerative disorder; gene therapy; growth deficiency;
KW cirrhosis; hypoproliferative disorder; lesion.
XX
OS Homo sapiens.
XX
PN WO200220032-A1.
XX
PD 14-MAR-2002.
XX
PF 10-SEP-2001; 2001WO-US028254.
XX
PR 08-SEP-2000; 2000US-0231212P.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
PA (UYSF-) UNIV SOUTH FLORIDA.
XX
PI Yu H, Pardoll D, Jove R, Dalton W;
XX
DR WPI; 2002-362218/39.
XX
PT Modulating angiogenesis and an immune response in an individual, for
PT treating a hypoxic or ischemic condition, comprises administering a
PT compound that modulates the activity of a signal transducer and activator
PT of transcription 3.
XX
PS Disclosure; Page 83-85; 94pp; English.
XX
CC The invention relates to a method of modulating angiogenesis and immune
CC response. Method involves administering to an individual a compound that
CC modulate the activity of signal transducer and activator of transcription
CC 3 (Stat3). Modulating angiogenesis is useful for treating or preventing
CC hypoxic or ischaemic condition or disorder which is the result of stroke,
CC ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,
CC tissue ischaemia in the lower extremities, infarction, trauma, vascular
CC occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,
CC chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,
CC epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest,
CC nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy
CC with neovascularisation. Suppressing an immune response is useful for
CC ameliorating a symptom of an autoimmune disease such as systemic lupus
CC erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,
CC Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,
CC mixed connective tissue disease, primary biliary cirrhosis, pernicious
CC anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,
CC gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,
CC idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's
CC disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune
CC infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and
CC dense deposit disease. The method is useful in preventing or treating
CC specific proliferative and oncogenic disease which includes sarcomas and
CC carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,
CC fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,
CC hypoproliferative disorders, physical trauma, lesions and wounds. The
CC method is also used in gene therapy. The present sequence is human
CC protein related to angiogenesis regulation
XX
SQ Sequence 769 AA;

Query Match 100.0%; Score 669; DB 5; Length 769;
Best Local Similarity 100.0%; Pred. No. 4.6e-68;
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377
```

RESULT 9
AAR72082
ID AAR72082 standard; protein; 770 AA.
XX
AC AAR72082;
XX
DT 25-MAR-2003 (revised)
DT 27-SEP-1995 (first entry)
XX
DE Mouse Stat3 (19sf6).
XX
KW Signal transducer and activator of transcription; STAT; 19sf6; Stat3;
KW receptor recognition factor; transcription factor; cellular debilitation;
KW derangement; dysfunction; interferon-gamma.
XX
OS Mus sp.
XX
PN WO9508629-A1.
XX
PD 30-MAR-1995.
XX
PF 26-SEP-1994; 94WO-US010849.
XX
PR 24-SEP-1993; 93US-00126588.
PR 24-SEP-1993; 93US-00126595.
PR 11-MAR-1994; 94US-00212184.
PR 11-MAR-1994; 94US-00212185.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI Darnell JE, Schindler CW, Shuai K, Wen Z, Zhong Z;
XX
DR WPI; 1995-139598/18.
DR N-PSDB; AAQ89340.
XX
PT Receptor recognition factor implicated in transcriptional stimulation of
PT genes - useful in drug screening assays and/or for treating cellular
PT debilitations, derangements and/or dysfunctions, etc.
XX
PS Claim 1; Page 107-110; 160pp; English.
XX
CC A fragment encoding the human Stat91 protein was used to screen a murine
CC thymus and spleen cDNA for homologous proteins. A highly homologous gene
CC (given in AAQ89338) was isolated that encoded a 91 kDa protein (AAR72080)
CC (Stat1) that was responsive to interferon- gamma. Using a fragment of the
CC mouse gene as probe, 2 additional members of the 113-91 family of
CC receptor recognition factor proteins were isolated. The 2 genes (AAQ89339
CC -40) were cloned in plasmids 13sf1 and 19sf6 and encoded proteins termed
CC Stat4 (AAR72081) and Stat3 (AAR72082), respectively. (Updated on 25-MAR-
CC 2003 to correct PN field.)
XX
SQ Sequence 770 AA;

```
  Query Match             100.0%;  Score 669;  DB 2;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-68;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377
```

RESULT 10
AAR82995
ID AAR82995 standard; protein; 770 AA.
XX
AC AAR82995;
XX
DT 25-MAR-1996 (first entry)
XX
DE Mouse liver acute phase response factor.
XX
KW Mouse; acute phase response factor; transcription factor; interleukin-6;
KW signal transmission; liver; antibody; antisense; ribozyme;
KW antiinflammatory; antitumor; hypotensive; therapy.
XX
OS Mus musculus.
XX
PN EP676469-A2.
XX
PD 11-OCT-1995.
XX
PF 29-MAR-1995; 95EP-00104670.
XX
PR 04-APR-1994; 94JP-00065825.
XX
PA (KISH/) KISHIMOTO T.
XX
PI Akira S, Kishimoto T;
XX
DR WPI; 1995-346089/45.
DR N-PSDB; AAT05619.
XX
PT New acute phase response factor - for developing inhibitory agents for
PT treating diseases induced by cytokine(s) such as IL-6, e.g. inflammatory
PT diseases.
XX
PS Claim 10; Page 20-22; 31pp; English.
XX
CC The sequence represents a mouse acute phase response factor (APRF), a
CC transcription factor related to signal transmission of interleukin-6 (IL-
CC 6). The protein is encoded by a cDNA, isolated from a mouse liver cDNA
CC library using a polymerase chain reaction product (amplified using
CC primers derived from an IL-6-treated mouse liver peptide) as a probe.
CC APRF-inhibitors, e.g. antibodies, antisense oligonucleotides or
CC ribozymes, may be used to treat diseases induced by IL-6, e.g.
CC inflammatory disease, leukemia, cancer, osteoclasia, pulmonary
CC hypertension, etc
XX
SQ Sequence 770 AA;

```
  Query Match             100.0%;  Score 669;  DB 2;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-68;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377
```

```
                 |||||||||
Db       369 DKDSGDVAA 377


RESULT 11
AAW03176
ID   AAW03176 standard; protein; 770 AA.
XX
AC   AAW03176;
XX
DT   24-OCT-1996  (first entry)
XX
DE   Mouse STAT4.
XX
KW   STAT; STAT4; signal transducer and activator of transcription;
KW   DNA binding protein; ligand; receptor; oncogenesis; inflammation;
KW   autoimmune disease; antagonist; therapy.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Domain          398. .508
FT                   /label= DNA_binding_domain
FT                   /note= "Claim 3, page 110"
XX
PN   WO9620954-A2.
XX
PD   11-JUL-1996.
XX
PF   28-DEC-1995;   95WO-US017025.
XX
PR   06-JAN-1995;   95US-00369796.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Darnell JE,  Wen Z,  Horvath CM,  Zhong Z;
XX
DR   WPI; 1996-333941/33.
DR   N-PSDB; AAT31280.
XX
PT   New STAT protein DNA-binding domain peptide(s) - useful for diagnosing,
PT   preventing or treating cellular dysfunction, e.g. oncogenesis,
PT   inflammation, parasitic disease or autoimmunity.
XX
PS   Disclosure; Page 87-90; 138pp; English.
XX
CC   Mouse signal transducer and activator of transcription (STAT) protein
CC   STAT4 (AAW03176) serves a dual purpose, i.e. signal transduction from
CC   ligand-activated receptor kinase complexes followed by nuclear
CC   translocation and DNA binding to activate transcription. Recombinant
CC   STAT4 can be obtd. using cDNA clone 19sf6 (AAT31278) obtd. from
CC   splenic/thymic cells. STAT4 includes a DNA-binding domain (see also
CC   AAW03167) capable of both receptor recognition and message delivery via
CC   DNA binding in a receptor-ligand specific manner. STAT proteins and their
CC   DNA binding domains (see also AAW03165-75) are useful for screening
CC   antagonists used to inhibit STAT-mediated signal transduction and
CC   activation of transcription
XX
SQ   Sequence 770 AA;

  Query Match             100.0%;  Score 669;  DB 2;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-68;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
```

```
                 |||||||||
Db       369 DKDSGDVAA 377


RESULT 12
AAY03768
ID   AAY03768 standard; protein; 770 AA.
XX
AC   AAY03768;
XX
DT   11-JUN-1999  (first entry)
XX
DE   Human STAT3 allelic variant.
XX
KW   Signal Transducer and Activator of Transcription 3; STAT3; allele; IL-6;
KW   intracellular transcription factor; interleukin-6; medicament; variant;
KW   pharmaceutical; autoimmune disease; inflammatory; human.
XX
OS   Homo sapiens.
XX
PN   EP905234-A2.
XX
PD   31-MAR-1999.
XX
PF   18-FEB-1998;   98EP-00102774.
XX
PR   16-SEP-1997;   97EP-00116061.
XX
PA   (ISTF ) ARS APPLIED RES SYSTEMS HOLDING NV.
XX
PI   Serlupi-Crescenzi O,  Della Pietra L;
XX
DR   WPI; 1999-192664/17.
DR   N-PSDB; AAX29281.
XX
PT   New human Signal Transducer and Activator of Transcription 3 (STAT3)
PT   allelic variant useful for treatment of autoimmune and inflammatory
PT   disease.
XX
PS   Claim 2; Page 9-13; 32pp; English.
XX
CC   The present sequence represents a predominant allelic variant of human
CC   Signal Transducer and Activator of Transcription 3 (STAT3) protein, an
CC   intracellular transcription factor which mediates IL-6 signals. The
CC   encoding sequence differs from the original published human STAT3 gene
CC   sequence (Akira et al., 1994). Host cells comprising a recombinant STAT3
CC   DNA molecule can be used for the recombinant expression of the variant.
CC   STAT3 protein is useful as a medicament or pharmaceutical composition for
CC   treatment of autoimmune or inflammatory diseases
XX
SQ   Sequence 770 AA;

  Query Match             100.0%;  Score 669;  DB 2;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-68;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
                 |||||||||
Db        369 DKDSGDVAA 377


RESULT 13
AAB12377
ID   AAB12377 standard; peptide; 770 AA.
XX
```

```
AC   AAB12377;
XX
DT   08-NOV-2000  (first entry)
XX
DE   N-terminal domain of murine STAT-3 protein.
XX
KW   STAT; signal transducer and activator of transcription; crystal;
KW   drug design; murine.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Region          4. .9
FT                   /label= Alpha helix 1
FT   Region          12. .21
FT                   /label= Alpha helix 2
FT   Region          19. .21
FT                   /label= 3(10) helix of alpha helix 2
FT   Region          28. .33
FT                   /label= Alpha helix 3
FT   Region          35. .40
FT                   /label= Alpha helix 4
FT   Region          43. .47
FT                   /label= Alpha helix 5
FT   Region          50. .73
FT                   /label= Alpha helix 6
FT   Region          77. .96
FT                   /label= Alpha helix 7
FT   Region          99. .119
FT                   /label= Alpha helix 8
XX
PN   US6087478-A.
XX
PD   11-JUL-2000.
XX
PF   23-JAN-1998;   98US-00012710.
XX
PR   23-JAN-1998;   98US-00012710.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Vinkemeier U,  Moarefi I,  Darnell JE,  Kuriyan J;
XX
DR   WPI; 2000-505108/45.
XX
PT   New crystals of an N-terminal fragment of a signal transducer and
PT   activator of transcription that effectively diffracts x-rays, useful for
PT   drug screening and development.
XX
PS   Disclosure; Fig 1; 42pp; English.
XX
CC   The present invention relates to a crystal of an N-terminal fragment of a
CC   signal transducer and activator of transcription (STAT) protein. The
CC   crystal effectively diffracts X-rays, allowing the determination of the
CC   atomic coordinates of the N-terminal domain to a resolution of greater
CC   than 5.0 Angstroms. The present sequence is the N-terminal domain of the
CC   murine STAT 3 protein. The N-terminal domain enables STAT dimers to
CC   interact and bind DNA cooperatively, a mechanism important for gene
CC   activation. The crystals are useful in drug screening and development by
CC   selecting a potential drug by performing rational drug design with the 3-
CC   dimensional structure determined for the crystal
XX
SQ   Sequence 770 AA;

  Query Match             100.0%;  Score 669;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-68;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


RESULT 14
AAE14652
ID   AAE14652 standard; protein; 770 AA.
XX
AC   AAE14652;
XX
DT   16-JUL-2002  (first entry)
XX
DE   Murine STAT3 protein.
XX
KW   Signal transducer and activator of transcription; STAT3;
KW   drug development; drug discovery; crystal; inflammation; allergy; asthma;
KW   leukaemia; anaemia; neutropaenia; thrombocytopaenia; cancer; obesity;
KW   viral disease; growth retardation; murine.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Domain          1. .130
FT                   /note= "Conserved N-terminal domain of the STAT family"
XX
PN   US6312887-B1.
XX
PD   06-NOV-2001.
XX
PF   24-APR-2000; 2000US-00556273.
XX
PR   23-JAN-1998;   98US-00012710.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Vinkemeier U,  Moarefi I,  Darnell JE,  Kuriyan J;
XX
DR   WPI; 2002-033337/04.
XX
PT   Identifying compounds that bind to signal transducer and activator of
PT   transcription proteins, useful for the production of new drugs.
XX
PS   Example; Col 47-50; 44pp; English.
XX
CC   The invention relates to methods for detecting compounds that bind to
CC   signal transducer and activator of transcription (STAT) proteins for the
CC   discovery and development of new drug compounds based on the structural
CC   properties of the protein crystal. The methods include: identifying a
CC   compound that binds to the N-terminal domain of a STAT protein,
CC   identifying a compound that enhances or diminishes the binding of the
CC   dimeric STAT proteins to each other and/or their nucleic acid binding
CC   site; or identifying a compound that enhances or diminishes the ability
CC   of STAT protein dimers to induce the expression of a gene operably under
CC   the control of a promoter containing at least two adjacent weak binding
CC   sites for STAT protein dimers. The methods are used for identifying new
CC   drugs. An antagonist of STAT N-terminal dimeric interactions that
CC   inhibits the binding of the STAT dimers to adjacent weak binding sites on
CC   a promoter of a gene, could be useful as drugs in the treatment of
CC   diseases, e.g. inflammation, allergy, asthma and leukaemias. On the other
CC   hand, an agonist of N-terminal dimeric interactions between STAT dimers,
CC   can be used as drugs in the treatment of diseases e.g. anaemia,
CC   neutropaenia, thrombocytopaenia, cancer, obesity, viral diseases and
CC   growth retardation. The present sequence is murine STAT3 protein
XX
SQ   Sequence 770 AA;

  Query Match             100.0%;  Score 669;  DB 5;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-68;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377
```

```
RESULT 15
ABG69497
ID   ABG69497 standard; protein; 770 AA.
XX
AC   ABG69497;
XX
DT   21-OCT-2002  (first entry)
XX
DE   Human bait protein STAT3.
XX
KW   Human; yeast two-hybrid assay; adipocyte; bait protein; NIDDM;
KW   non-insulin diabetes mellitus; obesity; selected interacting domain; SID;
KW   protein-protein interaction map; PIM; anorectic; metabolic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200253726-A2.
XX
PD   11-JUL-2002.
XX
PF   28-DEC-2001; 2001WO-EP015423.
XX
PR   02-JAN-2001; 2001US-0259377P.
XX
PA   (HYBR-) HYBRIGENICS.
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Legrain P,  Marullo S,  Jockers R;
XX
DR   WPI; 2002-583612/62.
DR   N-PSDB; ABS51033.
XX
PT   Novel complex of protein-protein interactions in adipocyte cells for
PT   identifying compounds that modulate the protein-protein interactions and
PT   useful for treating obesity and metabolic disorders.
XX
PS   Claim 1; Page 54; 125pp; English.
XX
CC   The invention relates to a complex of protein-protein interactions
CC   (forming a protein-protein interaction map, PIM) in adipocyte cells as
CC   defined in the specification, or polynucleotides in adipocytes encoding
CC   for the polypeptides. Also included are a recombinant cell expressing the
CC   interacting polypeptides and a method of selecting a modulating compound
CC   in adipocyte cells, by cultivating a recombinant host cell on a selective
CC   medium containing a modulating compound and a reporter gene the
CC   expression of which is toxic for the recombinant host cell which is
CC   transformed with two vectors, where the first vector comprises a
CC   polynucleotide encoding a first hybrid polypeptide and DNA binding domain
CC   and the second vector comprising a polynucleotide encoding a second
CC   hybrid polypeptide and an activating domain that activates the toxic
CC   reporter gene, when the first and second hybrid polypeptides interact and
CC   selecting the modulating compound which inhibits the growth of the
CC   recombinant host cell (i.e. using the yeast two-hybrid system). The
CC   complexes are useful for identifying compounds that modulate the protein-
CC   protein interactions and useful for treating obesity and metabolic
CC   disorders e.g. non-insulin dependent diabetes mellitus, NIDDM. The
CC   compound isolated by the method is useful for treating and preventing
CC   obesity or metabolic diseases. The interactions between the proteins of
CC   the complex further define a set of selected interacting domains, SID.
CC   The present sequence represents a member of the protein complex of the
CC   invention, used as the bait protein in the yeast two- hybrid assay
XX
SQ   Sequence 770 AA;

  Query Match            100.0%;  Score 669;  DB 5;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 4.6e-68;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377

Search completed: May 25, 2005, 17:36:36
Job time : 70.83 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:25:01 ; Search time 16.141 Seconds
(without alignments)
596.600 Million cell updates/sec

Title: US-10-090-185-24
Perfect score: 669
Sequence: 1 IACIGGPPNICLDRLENWIT.....LNYQLKIKVCIDKDSGDVAA 129

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	669	100.0	129	3	US-09-387-418A-24	Sequence 24, Appl
2	669	100.0	185	3	US-09-387-418A-23	Sequence 23, Appl
3	669	100.0	223	3	US-09-387-418A-22	Sequence 22, Appl
4	669	100.0	271	3	US-09-387-418A-9	Sequence 9, Appli
5	669	100.0	770	1	US-08-369-796-12	Sequence 12, Appl
6	669	100.0	770	1	US-08-416-581B-9	Sequence 9, Appli
7	669	100.0	770	2	US-08-852-091-12	Sequence 12, Appl
8	669	100.0	770	2	US-08-820-754-12	Sequence 12, Appl
9	669	100.0	770	3	US-08-956-652-12	Sequence 12, Appl
10	669	100.0	770	3	US-08-956-869-12	Sequence 12, Appl
11	669	100.0	770	3	US-09-012-710-8	Sequence 8, Appli
12	669	100.0	770	3	US-08-948-547-12	Sequence 12, Appl
13	669	100.0	770	3	US-09-364-970-3	Sequence 3, Appli
14	669	100.0	770	3	US-09-364-970-5	Sequence 5, Appli
15	669	100.0	770	3	US-09-556-273-8	Sequence 8, Appli
16	669	100.0	770	3	US-08-956-653A-12	Sequence 12, Appl
17	669	100.0	770	3	US-09-526-542-2	Sequence 2, Appli
18	669	100.0	770	4	US-08-212-185-12	Sequence 12, Appl
19	669	100.0	770	4	US-10-117-087-2	Sequence 2, Appli
20	669	100.0	770	5	PCT-US95-17025-12	Sequence 12, Appl
21	664	99.3	770	1	US-08-416-581B-1	Sequence 1, Appli
22	664	99.3	770	1	US-08-416-581B-5	Sequence 5, Appli
23	664	99.3	770	3	US-09-087-465-6	Sequence 6, Appli
24	664	99.3	770	4	US-09-972-800A-6	Sequence 6, Appli
25	664	99.3	771	1	US-08-276-099A-14	Sequence 14, Appl
26	664	99.3	771	1	US-08-781-890-14	Sequence 14, Appl
27	571	85.4	229	3	US-09-387-418A-18	Sequence 18, Appl
28	571	85.4	229	3	US-09-387-418A-28	Sequence 28, Appl
29	571	85.4	229	3	US-09-387-418A-30	Sequence 30, Appl
30	571	85.4	229	3	US-09-387-418A-31	Sequence 31, Appl
31	571	85.4	252	3	US-09-387-418A-14	Sequence 14, Appl
32	554	82.8	229	3	US-09-387-418A-29	Sequence 29, Appl
33	494	73.8	96	3	US-09-387-418A-25	Sequence 25, Appl
34	493	73.7	213	3	US-09-387-418A-19	Sequence 19, Appl
35	493	73.7	236	3	US-09-387-418A-15	Sequence 15, Appl
36	436	65.2	268	3	US-09-387-418A-12	Sequence 12, Appl
37	436	65.2	582	4	US-09-430-806A-3	Sequence 3, Appli
38	436	65.2	712	1	US-08-369-796-6	Sequence 6, Appli
39	436	65.2	712	2	US-08-852-091-6	Sequence 6, Appli
40	436	65.2	712	2	US-08-820-754-6	Sequence 6, Appli
41	436	65.2	712	3	US-08-956-652-6	Sequence 6, Appli
42	436	65.2	712	3	US-08-956-869-6	Sequence 6, Appli
43	436	65.2	712	3	US-08-948-547-6	Sequence 6, Appli
44	436	65.2	712	3	US-08-956-653A-6	Sequence 6, Appli
45	436	65.2	712	4	US-08-212-185-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-09-387-418A-24
; Sequence 24, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/09/387,418A
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-387-418A-24

Query Match 100.0%; Score 669; DB 3; Length 129;
Best Local Similarity 100.0%; Pred. No. 5.5e-69;
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120

Qy        121 DKDSGDVAA 129
              |||||||||
Db        121 DKDSGDVAA 129
```

RESULT 2
US-09-387-418A-23
; Sequence 23, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
; TITLE OF INVENTION: INTERACTIONS

```
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 23
;   LENGTH: 185
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-23

  Query Match             100.0%;  Score 669;  DB 3;  Length 185;
  Best Local Similarity   100.0%;  Pred. No. 8.9e-69;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         57 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 116

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        117 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 176

Qy        121 DKDSGDVAA 129
              |||||||||
Db        177 DKDSGDVAA 185


RESULT 3
US-09-387-418A-22
; Sequence 22, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
;  APPLICANT:  Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 22
;   LENGTH: 223
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-22

  Query Match             100.0%;  Score 669;  DB 3;  Length 223;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-68;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         95 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 154

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 214

Qy        121 DKDSGDVAA 129
              |||||||||
Db        215 DKDSGDVAA 223


RESULT 4
US-09-387-418A-9
; Sequence 9, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
;  APPLICANT:  Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 271
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-9

  Query Match             100.0%;  Score 669;  DB 3;  Length 271;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-68;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 202

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 262

Qy        121 DKDSGDVAA 129
              |||||||||
Db        263 DKDSGDVAA 271


RESULT 5
US-08-369-796-12
; Sequence 12, Application US/08369796
; Patent No. 5716622
;  GENERAL INFORMATION:
;    APPLICANT:  James E. Darnell, Jr.
;    APPLICANT:  Zilong Wen
;    APPLICANT:  Curt M. Horvath
;    APPLICANT:  Zhong Zhong
;    TITLE OF INVENTION:  FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
;    TITLE OF INVENTION:  TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/369,796
;      FILING DATE:  06-JAN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-116
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
```

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-369-796-12

  Query Match            100.0%;  Score 669;  DB 1;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 6e-68;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


RESULT 6
US-08-416-581B-9
; Sequence 9, Application US/08416581B
; Patent No. 5719042
;  GENERAL INFORMATION:
;    APPLICANT:  Kishimoto, Tadamitsu
;    APPLICANT:  Akira, Shizuo
;    TITLE OF INVENTION:  TRANSCRIPTION FACTOR APRF
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SUGHRUE, MION, ZINN, MACPEAK & SEAS
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/416,581B
;      FILING DATE:  04-APR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-65825/1994
;      FILING DATE:  04-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Nakamura, Dean H.
;      REGISTRATION NUMBER:  33,981
;      REFERENCE/DOCKET NUMBER:  Q-37891
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)293-7060
;      TELEFAX:  (202)293-7860
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-416-581B-9

  Query Match            100.0%;  Score 669;  DB 1;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 6e-68;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


RESULT 7
US-08-852-091-12
; Sequence 12, Application US/08852091
; Patent No. 5883228
;  GENERAL INFORMATION:
;    APPLICANT:  James E. Darnell, Jr.
;    APPLICANT:  Zilong Wen
;    APPLICANT:  Curt M. Horvath
;    APPLICANT:  Zhong Zhong
;    TITLE OF INVENTION:  FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
;    TITLE OF INVENTION:  TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/852,091
;      FILING DATE:  06-MAY-1997
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/369,796
;      FILING DATE:  06-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-116
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-852-091-12

  Query Match            100.0%;  Score 669;  DB 2;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 6e-68;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


RESULT 8
```

```
US-08-820-754-12
; Sequence 12, Application US/08820754
; Patent No. 5976835
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/820,754
;      FILING DATE:  19-MAR-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,185
;      FILING DATE:  11-MAR-1994
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-820-754-12

  Query Match            100.0%;  Score 669;  DB 2;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 6e-68;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377
```

```
RESULT 9
US-08-956-652-12
; Sequence 12, Application US/08956652
; Patent No. 6013475
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/956,652
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,185
;      FILING DATE:  11-MAR-1994
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-956-652-12

  Query Match            100.0%;  Score 669;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 6e-68;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368
```

```
Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


RESULT 10
US-08-956-869-12
; Sequence 12, Application US/08956869
; Patent No. 6030808
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/956,869
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/212,185
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-956-869-12

  Query Match             100.0%;  Score 669;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 6e-68;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


RESULT 11
US-09-012-710-8
; Sequence 8, Application US/09012710
; Patent No. 6087478
;  GENERAL INFORMATION:
;    APPLICANT:  Vinkemeier, Uwe
;    APPLICANT:  Moarefi, Ismail
;    APPLICANT:  Darnell, Jr., James E.
;    APPLICANT:  Kuriyan, John
;    TITLE OF INVENTION:  A CRYSTAL OF THE N-TERMINAL DOMAIN OF A
;    TITLE OF INVENTION:  STAT PROTEIN AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/012,710
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-194
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-09-012-710-8

  Query Match             100.0%;  Score 669;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 6e-68;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


RESULT 12
US-08-948-547-12
; Sequence 12, Application US/08948547
; Patent No. 6124118
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/948,547
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/212,185
;      FILING DATE:  11-MAR-1994
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-948-547-12

  Query Match             100.0%;  Score 669;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 6e-68;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


RESULT 13
US-09-364-970-3
; Sequence 3, Application US/09364970
; Patent No. 6235873
; GENERAL INFORMATION:
;  APPLICANT: Bromberg, Jacqueline
;  TITLE OF INVENTION: CONSTITUTIVELY ACTIVE TRANSCRIPTION FACTORS AND THEIR
;  TITLE OF INVENTION: USES FOR IDENTIFYING MODULATORS OF ACTIVITY INCLUDING
;  TITLE OF INVENTION: DYSPROLIFERATIVE CELLULAR CHANGES
;  FILE REFERENCE: 600-1-252
;  CURRENT APPLICATION NUMBER: US/09/364,970
;  CURRENT FILING DATE: 1999-07-31
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-364-970-3

  Query Match             100.0%;  Score 669;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 6e-68;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


RESULT 14
US-09-364-970-5
; Sequence 5, Application US/09364970
; Patent No. 6235873
; GENERAL INFORMATION:
;  APPLICANT: Bromberg, Jacqueline
;  TITLE OF INVENTION: CONSTITUTIVELY ACTIVE TRANSCRIPTION FACTORS AND THEIR
;  TITLE OF INVENTION: USES FOR IDENTIFYING MODULATORS OF ACTIVITY INCLUDING
;  TITLE OF INVENTION: DYSPROLIFERATIVE CELLULAR CHANGES
;  FILE REFERENCE: 600-1-252
;  CURRENT APPLICATION NUMBER: US/09/364,970
;  CURRENT FILING DATE: 1999-07-31
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-364-970-5

  Query Match             100.0%;  Score 669;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 6e-68;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377
```

```
RESULT 15
US-09-556-273-8
; Sequence 8, Application US/09556273
; Patent No. 6312887
;  GENERAL INFORMATION:
;    APPLICANT:  Vinkemeier, Uwe
;    APPLICANT:  Moarefi, Ismail
;    APPLICANT:  Darnell, Jr., James E.
;    APPLICANT:  Kuriyan, John
;    TITLE OF INVENTION:  A CRYSTAL OF THE N-TERMINAL DOMAIN OF A
;    TITLE OF INVENTION:  STAT PROTEIN AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/556,273
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/012,710
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-194
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-09-556-273-8

  Query Match           100.0%;  Score 669;  DB 3;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 6e-68;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


Search completed: May 25, 2005, 17:47:38
Job time : 17.141 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:36:58 ; Search time 56.4935 Seconds
(without alignments)
763.830 Million cell updates/sec

Title: US-10-090-185-24
Perfect score: 669
Sequence: 1 IACIGGPPNICLDRLENWIT.....LNYQLKIKVCIDKDSGDVAA 129

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1434725 seqs, 334507595 residues

Total number of hits satisfying chosen parameters: 1434725

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	669	100.0	129	13	US-10-090-185-24	Sequence 24, Appl
2	669	100.0	185	13	US-10-090-185-23	Sequence 23, Appl
3	669	100.0	223	13	US-10-090-185-22	Sequence 22, Appl
4	669	100.0	271	13	US-10-090-185-9	Sequence 9, Appli
5	669	100.0	720	15	US-10-380-020-4	Sequence 4, Appli
6	669	100.0	769	15	US-10-380-020-2	Sequence 2, Appli
7	669	100.0	769	15	US-10-380-020-5	Sequence 5, Appli
8	669	100.0	770	11	US-09-876-773-12	Sequence 12, Appl
9	669	100.0	770	14	US-10-045-792-8	Sequence 8, Appli
10	669	100.0	770	14	US-10-038-010-56	Sequence 56, Appl
11	669	100.0	770	14	US-10-117-087-2	Sequence 2, Appli
12	669	100.0	770	17	US-10-639-617-12	Sequence 12, Appl
13	669	100.0	793	9	US-09-925-302-780	Sequence 780, App
14	669	100.0	793	10	US-09-925-302-780	Sequence 780, App
15	666	99.6	770	15	US-10-116-275-329	Sequence 329, App
16	664	99.3	770	15	US-10-116-275-349	Sequence 349, App
17	571	85.4	229	13	US-10-090-185-18	Sequence 18, Appl
18	571	85.4	229	13	US-10-090-185-28	Sequence 28, Appl
19	571	85.4	229	13	US-10-090-185-30	Sequence 30, Appl
20	571	85.4	229	13	US-10-090-185-31	Sequence 31, Appl
21	571	85.4	252	13	US-10-090-185-14	Sequence 14, Appl
22	554	82.8	229	13	US-10-090-185-29	Sequence 29, Appl
23	494	73.8	96	13	US-10-090-185-25	Sequence 25, Appl
24	493	73.7	213	13	US-10-090-185-19	Sequence 19, Appl
25	493	73.7	236	13	US-10-090-185-15	Sequence 15, Appl
26	436	65.2	268	13	US-10-090-185-12	Sequence 12, Appl
27	436	65.2	582	14	US-10-245-120-3	Sequence 3, Appli
28	436	65.2	712	11	US-09-876-773-6	Sequence 6, Appli
29	436	65.2	712	14	US-10-245-120-2	Sequence 2, Appli
30	436	65.2	712	17	US-10-639-617-6	Sequence 6, Appli
31	436	65.2	712	17	US-10-936-390-5	Sequence 5, Appli
32	436	65.2	750	9	US-09-833-205-2	Sequence 2, Appli
33	436	65.2	750	11	US-09-876-773-4	Sequence 4, Appli
34	436	65.2	750	14	US-10-245-120-1	Sequence 1, Appli
35	436	65.2	750	14	US-10-308-279-44	Sequence 44, Appl
36	436	65.2	750	16	US-10-755-889-352	Sequence 352, App
37	436	65.2	750	16	US-10-755-889-823	Sequence 823, App
38	436	65.2	750	17	US-10-492-043-19	Sequence 19, Appl
39	436	65.2	750	17	US-10-639-617-4	Sequence 4, Appli
40	436	65.2	786	9	US-09-925-297-550	Sequence 550, App
41	432	64.6	749	9	US-09-833-205-4	Sequence 4, Appli
42	395	59.0	749	9	US-09-833-205-6	Sequence 6, Appli
43	395	59.0	749	11	US-09-876-773-8	Sequence 8, Appli
44	395	59.0	749	14	US-10-045-792-7	Sequence 7, Appli
45	395	59.0	749	14	US-10-205-194-39	Sequence 39, Appl

ALIGNMENTS

```
RESULT 1
US-10-090-185-24
; Sequence 24, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
;   LENGTH: 129
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-24

  Query Match           100.0%;  Score 669;  DB 13;  Length 129;
  Best Local Similarity  100.0%;  Pred. No. 9.6e-67;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
```

```
Qy        121 DKDSGDVAA 129
              |||||||||
Db        121 DKDSGDVAA 129


RESULT 2
US-10-090-185-23
; Sequence 23, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
;   LENGTH: 185
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-23

  Query Match            100.0%;  Score 669;  DB 13;  Length 185;
  Best Local Similarity  100.0%;  Pred. No. 1.5e-66;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         57 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 116

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        117 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 176

Qy        121 DKDSGDVAA 129
              |||||||||
Db        177 DKDSGDVAA 185


RESULT 3
US-10-090-185-22
; Sequence 22, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
;   LENGTH: 223
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-22

  Query Match            100.0%;  Score 669;  DB 13;  Length 223;
  Best Local Similarity  100.0%;  Pred. No. 1.9e-66;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         95 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 154

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 214

Qy        121 DKDSGDVAA 129
              |||||||||
Db        215 DKDSGDVAA 223


RESULT 4
US-10-090-185-9
; Sequence 9, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 271
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-9

  Query Match            100.0%;  Score 669;  DB 13;  Length 271;
  Best Local Similarity  100.0%;  Pred. No. 2.5e-66;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 202

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 262

Qy        121 DKDSGDVAA 129
              |||||||||
Db        263 DKDSGDVAA 271


RESULT 5
US-10-380-020-4
; Sequence 4, Application US/10380020
; Publication No. US20040052762A1
; GENERAL INFORMATION:
;  APPLICANT: Yu, Hua
;  APPLICANT:  Pardoll, Drew
;  APPLICANT:  Jove, Richard
;  APPLICANT:  Dalton, Willian
;  TITLE OF INVENTION: Stat3 Agonists and Antagonists and Therapeutic Uses Thereof
;  FILE REFERENCE: 10873-009-999
;  CURRENT APPLICATION NUMBER: US/10/380,020
;  CURRENT FILING DATE:  2003-03-07
;  PRIOR APPLICATION NUMBER: 60/231,212
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
```

```
;   LENGTH: 720
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-380-020-4

  Query Match            100.0%;  Score 669;  DB 15;  Length 720;
  Best Local Similarity  100.0%;  Pred. No. 8.8e-66;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


RESULT 6
US-10-380-020-2
; Sequence 2, Application US/10380020
; Publication No. US20040052762A1
; GENERAL INFORMATION:
;  APPLICANT: Yu, Hua
;  APPLICANT:  Pardoll, Drew
;  APPLICANT:  Jove, Richard
;  APPLICANT:  Dalton, Willian
;  TITLE OF INVENTION: Stat3 Agonists and Antagonists and Therapeutic Uses Thereof
;  FILE REFERENCE: 10873-009-999
;  CURRENT APPLICATION NUMBER: US/10/380,020
;  CURRENT FILING DATE:  2003-03-07
;  PRIOR APPLICATION NUMBER: 60/231,212
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 769
;   TYPE: PRT<
;   ORGANISM: Homo Sapiens
US-10-380-020-2

  Query Match            100.0%;  Score 669;  DB 15;  Length 769;
  Best Local Similarity  100.0%;  Pred. No. 9.6e-66;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


RESULT 7
US-10-380-020-5
; Sequence 5, Application US/10380020
; Publication No. US20040052762A1
; GENERAL INFORMATION:
;  APPLICANT: Yu, Hua
;  APPLICANT:  Pardoll, Drew
;  APPLICANT:  Jove, Richard
;  APPLICANT:  Dalton, Willian
;  TITLE OF INVENTION: Stat3 Agonists and Antagonists and Therapeutic Uses Thereof
;  FILE REFERENCE: 10873-009-999
;  CURRENT APPLICATION NUMBER: US/10/380,020
;  CURRENT FILING DATE:  2003-03-07
;  PRIOR APPLICATION NUMBER: 60/231,212
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 769
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-380-020-5

  Query Match            100.0%;  Score 669;  DB 15;  Length 769;
  Best Local Similarity  100.0%;  Pred. No. 9.6e-66;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


RESULT 8
US-09-876-773-12
; Sequence 12, Application US/09876773
; Publication No. US20040058318A1
;  GENERAL INFORMATION:
;       APPLICANT: Darnell Jr., James E.
;                  Schindler, Christian W.
;                  Fu, Xian-Yuan
;                  Wen, Zilong
;                  Zhong, Zhong
;       TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
;                           SEQUENCES AND METHODS OF USE THEREOF
;       NUMBER OF SEQUENCES: 25
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Klauber & Jackson
;            STREET: 411 Hackensack Avenue
;            CITY: Hackensack
;            STATE: New Jersey
;            COUNTRY: USA
;            ZIP: 07601
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/876,773
;            FILING DATE: 07-Jun-2001
;            CLASSIFICATION: 435
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/212,185
;            FILING DATE: 11-MAR-1994
;            APPLICATION NUMBER: US 07/980,498
;            FILING DATE: 23-NOV-1992
;            APPLICATION NUMBER: US 07/854,296
;            FILING DATE: 19-MAR-1992
;            APPLICATION NUMBER: WO US93/02569
;            FILING DATE: 19-MAR-1993
;            APPLICATION NUMBER: US 08/126,588
;            FILING DATE: 24-SEP-1993
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Jackson Esq., David A.
;            REGISTRATION NUMBER: 26,742
;            REFERENCE/DOCKET NUMBER: 600-1-073 CIP
```

```
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 201 487-5800
;            TELEFAX: 201 343-1684
;            TELEX: 133521
;   INFORMATION FOR SEQ ID NO: 12:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 770 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-876-773-12

  Query Match             100.0%;  Score 669;  DB 11;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 9.6e-66;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


RESULT 9
US-10-045-792-8
; Sequence 8, Application US/10045792
; Publication No. US20030003563A1
;   GENERAL INFORMATION:
;       APPLICANT: Vinkemeier, Uwe
;                  Moarefi, Ismail
;                  Darnell, Jr., James E.
;                  Kuriyan, John
;       TITLE OF INVENTION: A CRYSTAL OF THE N-TERMINAL DOMAIN OF A
;                           STAT PROTEIN AND METHODS OF USE THEREOF
;       NUMBER OF SEQUENCES: 13
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Klauber & Jackson
;            STREET: 411 Hackensack Avenue, 4th Floor
;            CITY: Hackensack
;            STATE: New Jersey
;            COUNTRY: USA
;            ZIP: 07601
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/045,792
;            FILING DATE: 19-Oct-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/012,710
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Jackson Esq., David A.
;            REGISTRATION NUMBER: 26,742
;            REFERENCE/DOCKET NUMBER: 600-1-194
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 201-487-5800
;            TELEFAX: 201-343-1684
;            TELEX: 133521
;   INFORMATION FOR SEQ ID NO: 8:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 770 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       HYPOTHETICAL: NO
;       SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-045-792-8

  Query Match             100.0%;  Score 669;  DB 14;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 9.6e-66;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


RESULT 10
US-10-038-010-56
; Sequence 56, Application US/10038010
; Publication No. US20030040089A1
; GENERAL INFORMATION:
;  APPLICANT: HYBRIGENICS
;  APPLICANT:  Pierre, Legrain
;  TITLE OF INVENTION: Protein-protein interactions in adipocyte cells
;  FILE REFERENCE: B4767A
;  CURRENT APPLICATION NUMBER: US/10/038,010
;  CURRENT FILING DATE:  2002-07-23
;  PRIOR APPLICATION NUMBER: US 60/259,377
;  PRIOR FILING DATE: 2001-01-02
;  NUMBER OF SEQ ID NOS: 67
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: STAT3 : Transcription factor
;   LOCATION: (1)..(770)
;   OTHER INFORMATION:
US-10-038-010-56

  Query Match             100.0%;  Score 669;  DB 14;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 9.6e-66;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


RESULT 11
US-10-117-087-2
; Sequence 2, Application US/10117087
; Publication No. US20030166854A1
; GENERAL INFORMATION:
;  APPLICANT: SERLUPI-CRESCENZI, Ottaviano
;  APPLICANT:  DELLA PIETRA, Linda
```

```
;  TITLE OF INVENTION: ALLELIC VARIANTS OF HUMAN STAT3
;  FILE REFERENCE: SERLUPI=2
;  CURRENT APPLICATION NUMBER: US/10/117,087
;  CURRENT FILING DATE:  2002-04-08
;  PRIOR APPLICATION NUMBER: US/09/526,542
;  PRIOR FILING DATE: 2000-03-19
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Human
US-10-117-087-2

  Query Match           100.0%;  Score 669;  DB 14;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 9.6e-66;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


RESULT 12
US-10-639-617-12
; Sequence 12, Application US/10639617
; Publication No. US20050079543A1
;   GENERAL INFORMATION:
;        APPLICANT: Darnell Jr., James E.
;                   Schindler, Christian W.
;                   Fu, Xian-Yuan
;                   Wen, Zilong
;                   Zhong, Zhong
;        TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
;                            SEQUENCES AND METHODS OF USE THEREOF
;        NUMBER OF SEQUENCES: 25
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Klauber & Jackson
;             STREET: 411 Hackensack Avenue
;             CITY: Hackensack
;             STATE: New Jersey
;             COUNTRY: USA
;             ZIP: 07601
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/639,617
;             FILING DATE: 12-Aug-2003
;             CLASSIFICATION: 435
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/212,185
;             FILING DATE: 11-MAR-1994
;             APPLICATION NUMBER: US 07/980,498
;             FILING DATE: 23-NOV-1992
;             APPLICATION NUMBER: US 07/854,296
;             FILING DATE: 19-MAR-1992
;             APPLICATION NUMBER: WO US93/02569
;             FILING DATE: 19-MAR-1993
;             APPLICATION NUMBER: US 08/126,588
;             FILING DATE: 24-SEP-1993
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Jackson Esq., David A.
;             REGISTRATION NUMBER: 26,742
;             REFERENCE/DOCKET NUMBER: 600-1-073 CIP
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 201 487-5800
;             TELEFAX: 201 343-1684
;             TELEX: 133521
;   INFORMATION FOR SEQ ID NO: 12:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 770 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-639-617-12

  Query Match           100.0%;  Score 669;  DB 17;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 9.6e-66;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


RESULT 13
US-09-925-302-780
; Sequence 780, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA104
;  CURRENT APPLICATION NUMBER: US/09/925,302
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05918
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 896
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 780
;   LENGTH: 793
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-925-302-780

  Query Match           100.0%;  Score 669;  DB 9;  Length 793;
  Best Local Similarity  100.0%;  Pred. No. 1e-65;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 331

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        332 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 391

Qy        121 DKDSGDVAA 129
              |||||||||
Db        392 DKDSGDVAA 400


RESULT 14
US-09-925-302-780
```

```
; Sequence 780, Application US/09925302
; Publication No. US20030064072A9
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA104
;  CURRENT APPLICATION NUMBER: US/09/925,302
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05918
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 896
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 780
;   LENGTH: 793
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-925-302-780

  Query Match             100.0%;  Score 669;  DB 10;  Length 793;
  Best Local Similarity   100.0%;  Pred. No. 1e-65;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 331

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        332 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 391

Qy        121 DKDSGDVAA 129
              |||||||||
Db        392 DKDSGDVAA 400


RESULT 15
US-10-116-275-329
; Sequence 329, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
;  APPLICANT: Elan Pharmaceutical Technology
;  APPLICANT:  O'Mahony, Daniel J.
;  APPLICANT:  Brayden, David
;  APPLICANT:  Byrne, Daragh
;  APPLICANT:  Lambkin, Imelda
;  APPLICANT:  Higgins, Lisa
;  TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
;  TITLE OF INVENTION:  Compositions Targeting Peyer's Patches and M Cell Receptors
;  FILE REFERENCE: E1067/20087
;  CURRENT APPLICATION NUMBER: US/10/116,275
;  CURRENT FILING DATE:  2002-10-04
;  NUMBER OF SEQ ID NOS: 349
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 329
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Rat
US-10-116-275-329

  Query Match             99.6%;  Score 666;  DB 15;  Length 770;
  Best Local Similarity   99.2%;  Pred. No. 2.1e-65;
  Matches  128;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377

Search completed: May 25, 2005, 18:21:51
Job time : 59.8269 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:24:07 ; Search time 12.1698 Seconds
(without alignments)
1019.898 Million cell updates/sec

Title: US-10-090-185-24
Perfect score: 669
Sequence: 1 IACIGGPPNICLDRLENWIT.....LNYQLKIKVCIDKDSGDVAA 129

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	669	100.0	770	2	I49508	ISGF3 p91-related
2	664	99.3	770	2	A54444	DNA-binding protei
3	390.5	58.4	739	2	A46159	interferon-depende
4	375	56.1	748	2	A56047	gamma-interferon a
5	257	38.4	851	2	A46160	interferon alpha-i
6	160.5	24.0	786	2	I49274	mammary gland fact
7	160.5	24.0	793	2	S54772	mammary gland fact
8	160.5	24.0	794	2	G02317	transcription acti
9	142	21.2	794	2	S55527	mammary gland fact
10	98	14.6	848	2	A54740	interleukin-4-indu
11	93.5	14.0	837	2	I57557	DNA-Binding Protei
12	80.5	12.0	464	2	H90279	microtubule bindin
13	80	12.0	393	2	T25063	hypothetical prote
14	79	11.8	589	1	NDECAS	type I site-specif
15	78.5	11.7	191	2	B70042	hypothetical prote
16	78.5	11.7	253	2	F84840	hypothetical prote
17	78	11.7	498	2	T06667	argininosuccinate
18	77	11.5	590	2	A48461	ovarian abundant m
19	75.5	11.3	257	2	B96003	probable transcrip
20	75.5	11.3	926	2	T39082	hypothetical prote
21	75	11.2	121	2	G90550	hypothetical prote
22	75	11.2	623	2	G81420	hypothetical prote
23	75	11.2	695	2	T39954	probable gtpase ac
24	74.5	11.1	256	2	T48076	60S RIBOSOMAL PROT
25	74.5	11.1	900	2	T14277	myosin-like protei
26	74.5	11.1	2783	1	A41948	alpha-fetoprotein
27	74	11.1	893	2	F72253	hypothetical prote
28	73	10.9	546	2	AC2368	hypothetical prote
29	73	10.9	575	2	AF2374	hypothetical prote
30	72.5	10.8	257	2	T00423	ribosomal protein
31	72.5	10.8	432	2	A85040	hypothetical prote
32	72.5	10.8	447	2	S54631	HST3 protein - yea
33	72.5	10.8	1218	2	T14265	golgin-245 - mouse
34	72.5	10.8	1331	1	XORTDH	xanthine dehydroge
35	72	10.8	335	2	T25110	hypothetical prote
36	72	10.8	346	2	F87844	protein T22C1.6 [i
37	72	10.8	393	2	S32458	4-hydroxyphenylpyr
38	72	10.8	393	2	S32821	4-hydroxyphenylpyr
39	72	10.8	1999	1	S21801	myosin heavy chain
40	71.5	10.7	217	1	S73360	heat shock protein
41	71.5	10.7	328	2	A99345	conserved hypothet
42	71.5	10.7	573	2	A86253	hypothetical prote
43	71.5	10.7	879	2	F96558	probable protein k
44	71.5	10.7	1529	2	A59310	unconventional myo
45	71.5	10.7	3329	2	T42205	breast cancer susc

ALIGNMENTS

RESULT 1
I49508
ISGF3 p91-related transcription factor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I49508; I49009
R;Akira, S.; Nishio, Y.; Inoue, M.; Wang, X.J.; Wei, S.; Matsusaka, T.; Yoshida, K.; Su
Cell 77, 63-71, 1994
A;Title: Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related tr
A;Reference number: A54444; MUID:94208062; PMID:7512451
A;Accession: I49508
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-770 <RES>
A;Cross-references: UNIPROT:P42227; GB:L29278; NID:g476715; PIDN:AAA37254.1; PID:g47671
R;Raz, R.; Durbin, J.E.; Levy, D.E.
J. Biol. Chem. 269, 24391-24395, 1994
A;Title: Acute phase response factor and additional members of the interferon-stimulate
A;Reference number: I49009; MUID:95014185; PMID:7523373
A;Accession: I49009
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-393,'M',395-700,702-770 <RE2>
A;Cross-references: EMBL:U08378; NID:g473889; PIDN:AAA56668.1; PID:g473890
C;Genetics:
A;Gene: APRF
C;Superfamily: human signal transducer and transcription activator STAT5A

Query Match 100.0%; Score 669; DB 2; Length 770;
Best Local Similarity 100.0%; Pred. No. 1.8e-55;
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377
```

RESULT 2
A54444
DNA-binding protein APRF - human
C;Species: Homo sapiens (man)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 09-Jul-2004
C;Accession: A54444
R;Akira, S.; Nishio, Y.; Inoue, M.; Wang, X.J.; Wei, S.; Matsusaka, T.; Yoshida, K.; Su

```
Cell 77, 63-71, 1994
A;Title: Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related tra
A;Reference number: A54444; MUID:94208062; PMID:7512451
A;Accession: A54444
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-770 <RES>
A;Cross-references: UNIPROT:P40763; GB:L29277; NID:g475788; PID:g475789
C;Genetics:
A;Gene: GDB:STAT3; APRF
A;Cross-references: GDB:358950
A;Map position: 17q21-17q21
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: DNA binding; transcription factor

  Query Match             99.3%;  Score 664;  DB 2;  Length 770;
  Best Local Similarity   99.2%;  Pred. No. 5.5e-55;
  Matches  128;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELHQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


RESULT 3
A46159
interferon-dependent positive-acting transcription factor ISGF-3 91K chain - human
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A46159
R;Schindler, C.; Fu, X.Y.; Improta, T.; Aebersold, R.; Darnell Jr., J.E.
Proc. Natl. Acad. Sci. U.S.A. 89, 7836-7839, 1992
A;Title: Proteins of transcription factor ISGF-3: one gene encodes the 91-and 84-kDa ISG
A;Reference number: A46159; MUID:92366557; PMID:1502203
A;Accession: A46159
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid; protein
A;Residues: 1-739 <SCH>
A;Cross-references: UNIPROT:P42224
A;Experimental source: HeLa cells
A;Note: sequence extracted from NCBI backbone (NCBIP:110818)
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             58.4%;  Score 390.5;  DB 2;  Length 739;
  Best Local Similarity   60.7%;  Pred. No. 4.3e-29;
  Matches   74;  Conservative   18;  Mismatches   19;  Indels   11;  Gaps    1;

Qy          2 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 61
              |||||||| |||:|:           | |||:||||||:|| :|: ||| ::: :| :|
Db        246 ACIGGPPNACLDQLQ----------QVRQQLKKLEELEQKYTYEHDPITKNKQVLWDRT 294

Qy         62 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCID 121
                ||: |::|:|||||||||| || ||||:||||||| |:|||||  |||| ||:||  |
Db        295 FSLFQQLIQSSFVVERQPCMPTHPQRPLVLKTGVQFTVKLRLLVKLQELNYNLKVKVLFD 354

Qy        122 KD 123
              ||
Db        355 KD 356


RESULT 4
A56047
gamma-interferon activation site-binding protein Stat4 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004
C;Accession: A56047
R;Yamamoto, K.; Quelle, F.W.; Thierfelder, W.E.; Kreider, B.L.; Gilbert, D.J.; Jenkins,
Mol. Cell. Biol. 14, 4342-4349, 1994
A;Title: Stat4, a novel gamma interferon activation site-binding protein expressed in e
A;Reference number: A56047; MUID:94277038; PMID:8007943
A;Accession: A56047
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-748 <YAM>
A;Cross-references: UNIPROT:P42228; GB:U09351; NID:g509502; PIDN:AAA19692.1; PID:g50950
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: DNA binding; phosphoprotein

  Query Match             56.1%;  Score 375;  DB 2;  Length 748;
  Best Local Similarity   59.3%;  Pred. No. 1.3e-27;
  Matches   73;  Conservative   19;  Mismatches   31;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||| :  ||:|:|  | ||||  | |||::||:|   |::|:||||   | | ||
Db        244 IACIGGPLHNGLDQLQNCFTLLAESLFQLRQQLEKLQEQSTKMTYEGDPIPAQRAHLLER 303

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
                |  || |::|||||||||| || ||:|:|| :||| |:|||:| ||||||:|:|  |
Db        304 ATFLIYNLFKNSFVVERQPCMPTHPQRPMVLKTLIQFTVKLRLLIKLPELNYQVKVKASI 363

Qy        121 DKD 123
              |||
Db        364 DKN 366


RESULT 5
A46160
interferon alpha-induced transcription activator ISGF-3, 113K chain - human
N;Alternate names: stat2 protein
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A46160; S71908; S53873
R;Fu, X.Y.; Schindler, C.; Improta, T.; Aebersold, R.; Darnell Jr., J.E.
Proc. Natl. Acad. Sci. U.S.A. 89, 7840-7843, 1992
A;Title: The proteins of ISGF-3, the interferon alpha-induced transcriptional activator
A;Reference number: A46160; MUID:92366558; PMID:1502204
A;Accession: A46160
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA; protein
A;Residues: 1-851 <FU1>
A;Cross-references: UNIPROT:P52630
A;Note: sequence extracted from NCBI backbone (NCBIP:110820)
R;Yan, R.; Qureshi, S.; Zhong, Z.; Wen, Z.; Darnell, J.E.
submitted to the EMBL Data Library, December 1994
A;Reference number: S71908
A;Accession: S71908
A;Molecule type: DNA
A;Residues: 1-851 <YAN>
A;Cross-references: EMBL:U18671; NID:g1293919; PIDN:AAA98760.1; PID:g1293920
R;Yan, R.; Qureshi, S.; Zhong, Z.; Wen, Z.; Darnell Jr., J.E.
Nucleic Acids Res. 23, 459-463, 1995
A;Title: The genomic structure of the STAT genes: multiple exons in coincident sites in
A;Reference number: S53873; MUID:95192056; PMID:7885841
A;Accession: S53873
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-196;392-591;684-730 <YAW>
A;Cross-references: EMBL:U18671
C;Genetics:
A;Gene: stat2
A;Introns: 44/2; 95/3; 127/3; 157/3; 183/1; 211/3; 261/2; 314/2; 345/2; 365/2; 372/2; 4
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: signal transduction; transcription regulation

  Query Match             38.4%;  Score 257;  DB 2;  Length 851;
  Best Local Similarity   42.6%;  Pred. No. 2.3e-16;
  Matches   52;  Conservative   25;  Mismatches   45;  Indels    0;  Gaps    0;
```

```
Qy          2 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 61
              |||  | :  |::|| | |: |:     || :|:|: |   |||: ||: :   :   ::
Db        245 ACIRAPIDHGLEQLETWFTAGAKLLFHLRQLLKELKGLSCLVSYQDDPLTKGVDLRNAQV 304

Qy         62 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCID 121
               || : |:  ||||| |||||  | |||::||| :|| : ||||:  | |  | ::| ||
Db        305 TELLQRLLHRAFVVETQPCMPQTPHRPLILKTGSKFTVRTRLLVRLQEGNESLTVEVSID 364

Qy        122 KD 123
              ::
Db        365 RN 366


RESULT 6
I49274
mammary gland factor - mouse
N;Alternate names: STAT5 protein homolog p80
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I49274; S54773; S54727
R;Liu, X.; Robinson, G.W.; Gouilleux, F.; Groner, B.; Hennighausen, L.
Proc. Natl. Acad. Sci. U.S.A. 92, 8831-8835, 1995
A;Title: Cloning and expression of Stat5 and an additional homologue (Stat5b) involved i
A;Reference number: I49273; MUID:96004632; PMID:7568026
A;Accession: I49274
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-786 <RES>
A;Cross-references: UNIPROT:P42232; UNIPROT:Q9JKM1; EMBL:U21110; NID:g747973; PIDN:AAC52
R;Mui, A.L.F.; Wakao, H.; O'Farrell, A.M.; Harada, N.; Miyajima, A.
EMBO J. 14, 1166-1175, 1995
A;Title: Interleukin-3, granulocyte-macrophage colony stimulating factor and interleukin
A;Reference number: S54772; MUID:95237198; PMID:7720707
A;Accession: S54773
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-432,'E',434-786 <MUI>
A;Cross-references: EMBL:Z48539; NID:g758635; PIDN:CAA88420.1; PID:g758636
R;Azam, M.; Erdjument-Bromage, H.; Kreider, B.L.; Xia, M.; Quelle, F.; Basu, R.; Saris,
EMBO J. 14, 1402-1411, 1995
A;Title: Interleukin-3 signals through multiple isoforms of Stat5.
A;Reference number: S54725; MUID:95246733; PMID:7537213
A;Accession: S54727
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-432,'E',434-786 <AZA>
C;Genetics:
A;Gene: Stat5b
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             24.0%;  Score 160.5;  DB 2;  Length 786;
  Best Local Similarity   37.6%;  Pred. No. 2.9e-07;
  Matches   38;  Conservative   17;  Mismatches   37;  Indels    9;  Gaps    2;

Qy          5 GGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERIVEL 64
              ||||   || |::|   |||   | ||||:: | | |::   | |: :    :   | ::
Db        265 GGPPEGSLDVLQSWCEKLAEIIWQNRQQIRRAEHLCQQLPIPG-PVEEMLAEVNATITDI 323

Qy         65 FRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLV 105
                |: | |::|:||        | |:||  :|   |||||
Db        324 ISALVTSTFIIEKQP--------PQVLKTQTKFAATVRLLV 356


RESULT 7
S54772
mammary gland factor - mouse
N;Alternate names: stat5 protein
C;Species: Mus musculus (house mouse)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S54772; I49273
R;Mui, A.L.F.; Wakao, H.; O'Farrell, A.M.; Harada, N.; Miyajima, A.
EMBO J. 14, 1166-1175, 1995
A;Title: Interleukin-3, granulocyte-macrophage colony stimulating factor and interleukin
A;Reference number: S54772; MUID:95237198; PMID:7720707
A;Accession: S54772
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-793 <MUI>
A;Cross-references: UNIPROT:P42230; UNIPROT:Q9JIA0; EMBL:Z48538; NID:g758633; PIDN:CAA8
R;Liu, X.; Robinson, G.W.; Gouilleux, F.; Groner, B.; Hennighausen, L.
Proc. Natl. Acad. Sci. U.S.A. 92, 8831-8835, 1995
A;Title: Cloning and expression of Stat5 and an additional homologue (Stat5b) involved
A;Reference number: I49273; MUID:96004632; PMID:7568026
A;Accession: I49273
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-793 <RES>
A;Cross-references: EMBL:U21103; NID:g747971; PIDN:AAA80590.1; PID:g747972
C;Genetics:
A;Gene: Stat5a
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             24.0%;  Score 160.5;  DB 2;  Length 793;
  Best Local Similarity   37.6%;  Pred. No. 3e-07;
  Matches   38;  Conservative   17;  Mismatches   37;  Indels    9;  Gaps    2;

Qy          5 GGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERIVEL 64
              ||||   || |::|   |||   | ||||:: | | |::   | |: :    :   | ::
Db        265 GGPPEGSLDVLQSWCEKLAEIIWQNRQQIRRAEHLCQQLPIPG-PVEEMLAEVNATITDI 323

Qy         65 FRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLV 105
                |: | |::|:||        | |:||  :|   |||||
Db        324 ISALVTSTFIIEKQP--------PQVLKTQTKFAATVRLLV 356


RESULT 8
G02317
transcription activator stat5A - human
C;Species: Homo sapiens (man)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 09-Jul-2004
C;Accession: G02317
R;Lin, J.
submitted to the EMBL Data Library, December 1995
A;Reference number: H01043
A;Accession: G02317
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-794 <LIN>
A;Cross-references: UNIPROT:P42229; EMBL:U43185; NID:g1151169; PIDN:AAB06589.1; PID:g11
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             24.0%;  Score 160.5;  DB 2;  Length 794;
  Best Local Similarity   37.6%;  Pred. No. 3e-07;
  Matches   38;  Conservative   17;  Mismatches   37;  Indels    9;  Gaps    2;

Qy          5 GGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERIVEL 64
              ||||   || |::|   |||   | ||||:: | | |::   | |: :    :   | ::
Db        265 GGPPEGSLDVLQSWCEKLAEIIWQNRQQIRRAEHLCQQLPIPG-PVEEMLAEVNATITDI 323

Qy         65 FRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLV 105
                |: | |::|:||        | |:||  :|   |||||
Db        324 ISALVTSTFIIEKQP--------PQVLKTQTKFAATVRLLV 356


RESULT 9
S55527
mammary gland factor - sheep
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S55527; S44353
R;Wakao, H.; Gouilleux, F.; Groner, B.
EMBO J. 14, 854-855, 1995
A;Title: Corrigenda. Mammary gland factor (MGF) is a novel member of the cytokine regul
A;Reference number: S55527; MUID:95188889; PMID:7882987
A;Accession: S55527
A;Molecule type: mRNA
```

```
A;Residues: 1-794 <WAK>
A;Cross-references: UNIPROT:P42231; EMBL:X78428; NID:g602354; PIDN:CAA55191.1; PID:g6023
A;Note: this is a revision to the sequence from reference S44353
R;Wakao, H.; Gouilleux, F.; Groner, B.
EMBO J. 13, 2182-2191, 1994
A;Title: Mammary gland factor (MGF) is a novel member of the cytokine regulated transcri
A;Reference number: S44353; MUID:94244619; PMID:7514531
A;Accession: S44353
A;Molecule type: mRNA
A;Residues: 1-716,'RHLHGPGSLPSR',729,'P',731,'ASL' <WAW>
A;Cross-references: EMBL:X78428
A;Note: this sequence has been revised in reference S55527
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             21.2%;  Score 142;  DB 2;  Length 794;
  Best Local Similarity   36.4%;  Pred. No. 1.7e-05;
  Matches   36;  Conservative   17;  Mismatches   36;  Indels   10;  Gaps    3;

Qy          7 PPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERIVELFR 66
              ||   || |::|   |||   | ||||:: | | |::   | |: :    :   | ::
Db        268 PPR-SLDVLQSWCEKLAEIIWQNRQQIRRAEHLCQQLPIPG-PVEEMLAEVNATITDIIS 325

Qy         67 NLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLV 105
                |: | |::|:||       | |:||  :|   |||||
Db        326 ALVTSTFIIEKQP--------PQVLKTQTKFAATVRLLV 356


RESULT 10
A54740
interleukin-4-induced transcription factor stat - human
C;Species: Homo sapiens (man)
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004
C;Accession: A54740
R;Hou, J.; Schindler, U.; Henzel, W.J.; Ho, T.C.; Brasseur, M.; McKnight, S.L.
Science 265, 1701-1706, 1994
A;Title: An interleukin-4-induced transcription factor: IL-4 stat.
A;Reference number: A54740; MUID:94367369; PMID:8085155
A;Accession: A54740
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-848 <HOU>
A;Cross-references: UNIPROT:P42226
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: DNA binding; transcription regulation

  Query Match             14.6%;  Score 98;  DB 2;  Length 848;
  Best Local Similarity   29.1%;  Pred. No. 0.27;
  Matches   32;  Conservative   17;  Mismatches   39;  Indels   22;  Gaps    3;

Qy         10 ICLDRLENW-------------ITSLAESQLQTRQQIKKLEELQQKVSYKGDPI-VQHRP 55
              : | |:: |                |||  | :    :    :|||:|   |  :  : |
Db        196 LVLKRIQIWKRQQQLAGNGAPFEESLAPLQERCESLVDIYSQLQQEVGAAGGELEPKTRA 255

Qy         56 MLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLV 105
                |  |: |: | |: | |:||:||        | |:||  :|   || |:
Db        256 SLTGRLDEVLRTLVTSCFLVEKQP--------PQVLKTQTKFQAGVRFLL 297


RESULT 11
I57557
DNA-Binding Protein and transcription factor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C;Accession: I57557
R;Quelle, F.W.; Shimoda, K.; Thierfelder, W.; Fischer, C.L.; Kim, A.; Ruben, S.M.; Cleve
Mol. Cell. Biol. 15, 3336-3343, 1995
A;Title: Cloning of murine Stat6 and human Stat6, Stat proteins that are tyrosine phosph
A;Reference number: I57557; MUID:95280934; PMID:7760829
A;Accession: I57557
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-837 <RES>
A;Cross-references: UNIPROT:P52633; GB:L47650; NID:g1008876; PIDN:AAA79006.1; PID:g10088
C;Genetics:
A;Gene: STAT6
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             14.0%;  Score 93.5;  DB 2;  Length 837;
  Best Local Similarity   32.4%;  Pred. No. 0.71;
  Matches   33;  Conservative   12;  Mismatches   38;  Indels   19;  Gaps    3;

Qy          5 GGPPNICLDRLENWITSLAESQLQTRQQIKKLE-ELQQKVSYKGDPIVQHRPMLEERIVE 63
              | |    |  |:    || |   |  |:|    ||: |           |  |  |: |
Db        214 GTPFEESLAGLQERCESLVEIYSQLHQEIGAASGELEPKT----------RASLISRLDE 263

Qy         64 LFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLV 105
              : | |: |:|:||:||        | |:||  :|   || |:
Db        264 VLRTLVTSSFLVEKQP--------PQVLKTQTKFQAGVRFLL 297


RESULT 12
H90279
microtubule binding protein, probable [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: H90279
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder,
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: H90279
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-464 <KUR>
A;Cross-references: UNIPROT:Q9UXN4; GB:AE006641; NID:g13814451; PIDN:AAK41495.1; GSPDB:
C;Genetics:
A;Gene: SSO1256

  Query Match             12.0%;  Score 80.5;  DB 2;  Length 464;
  Best Local Similarity   27.9%;  Pred. No. 6;
  Matches   39;  Conservative   25;  Mismatches   41;  Indels   35;  Gaps    7;

Qy         14 RLENWITSLAESQLQTRQQIKKLEELQQKV--SYKG--DPIVQHRPML----------E 58
              :||| :  | |:| :| ::| ||||:  |:  |  |  : | :  |  |          |
Db        280 KLENAVEQLVEAQKRTDERITKLEEVTMKLVESQLGMQNEIRELRKALGSMGKRWGRDFE 339

Qy         59 ERIVELFRNLMK---------SAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPE 109
              : |:|:   | |         : |  :         |  |    ||::   | ||:|  :
Db        340 KLIIEIVDELAKQEGLDLKYVNKFTYK--------DDNGLFGLKGVEY--DVDLLIKDTK 389

Qy        110 LNYQLKIKVCIDKDSGDVAA 129
              : | ::||  ::||  : ||
Db        390 V-YLIEIKSYVEKDDVNWAA 408


RESULT 13
T25063
hypothetical protein T21C12.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 12-Jul-2004
C;Accession: T25063
R;Wilkinson, J.
submitted to the EMBL Data Library, July 1995
A;Reference number: Z19976
A;Accession: T25063
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-393 <WIL>
A;Cross-references: UNIPROT:Q22633; EMBL:Z50016; PIDN:CAA90315.1; GSPDB:GN00021; CESP:T
A;Experimental source: clone T21C12
C;Genetics:
A;Gene: CESP:T21C12.2
A;Map position: 3
A;Introns: 137/3; 236/3
```

```
C;Superfamily: 4-hydroxyphenylpyruvate_dioxygenase

  Query Match             12.0%;  Score 80;  DB 2;  Length 393;
  Best Local Similarity   20.8%;  Pred. No. 5.5;
  Matches   35;  Conservative   23;  Mismatches   56;  Indels   54;  Gaps    6;

Qy          3 CIGGPPNICLDRLENW---------ITSLAESQLQTRQQI------------------- 33
              |:|  |:: :|    |           |:·:| : |
Db        183 CVGNQPDLQMDSAVQWYEKVLKFHRFWSVDDSMIHTEYSALRSIVVTNFEETIKMPINEP 242

Qy         34 -----KKLEELQQKVSYKGDPIVQHRPMLEERIVEL----------FRNLMKSAF--VVE 76
                 | : ::|: | | |   |||  :    |:           | ::  | :  : |
Db        243 ATSDKKAISQIQEYVDYYGGSGVQHIALNTSDIITAIEALRARGCEFLSIPSSYYDNLKE 302

Qy         77 RQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKI--KVCIDK 122
              |      |      |:|  :    |: :|| | |  | |:|  | | |:
Db        303 RLAASSM------VVKEDMDRLQKLHILVDFDENGYLLQIFSKPCQDR 344


RESULT 14
NDECAS
type I site-specific deoxyribonuclease (EC 3.1.21.3) EcoA chain S - Escherichia coli
N;Alternate names: deoxyribonuclease, EcoA, S chain (ATP- and S-adenosyl-L-methionine-de
C;Species: Escherichia coli
C;Date: 30-Sep-1990 #sequence_revision 30-Sep-1990 #text_change 30-Jun-1993
C;Accession: A32343
R;Cowan, G.M.; Gann, A.A.F.; Murray, N.E.
Cell 56, 103-109, 1989
A;Title: Conservation of complex DNA recognition domains between families of restriction
A;Reference number: A32343; MUID:89089749; PMID:2642743
A;Accession: A32343
A;Molecule type: DNA
A;Residues: 1-589 <COW>
C;Comment: This S chain, one of three components (S, R, and M chains) of type I site-spe
confer specificity for the target sequence, GAG.
C;Superfamily: type I site-specific deoxyribonuclease EcoA chain S
C;Keywords: DNA binding; hydrolase; restriction modification system

  Query Match             11.8%;  Score 79;  DB 1;  Length 589;
  Best Local Similarity   24.4%;  Pred. No. 11;
  Matches   31;  Conservative   25;  Mismatches   41;  Indels   30;  Gaps    6;

Qy         10 ICLDRLENWITSLAESQLQTRQQ-------------IKKLEELQQKVSYKG-----DPIV 51
              : :| :| |      | ||||                |||| ||  ::: :|     ||
Db          6 LIVDHMETW-----TSALQTRSTAGRGSSGKIDLYGIKKLRELILELAVRGKLVPQDPND 60

Qy         52 QHRPMLEERIVELFRNLMKSAFVVERQPCMP--MHPDRPLVIKTGVQFTTKVRLLVKFPE 109
              :    | :||      |:|   : :::| :|     ::|  :  | ::||    | :  |
Db         61 EPASELLKRIAAEKAELVKQGKIKKQKP-LPEISEEEKPFELPDGWEWTT----LTRIAE 115

Qy        110 LNYQLKI 116
              :| :: :
Db        116 INPKIDV 122


RESULT 15
B70042
hypothetical protein yviE - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: B70042
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: B70042
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-191 <KUN>
A;Cross-references: UNIPROT:P96502; GB:Z99122; GB:AL009126; NID:g2636029; PIDN:CAB15556
A;Experimental source: strain 168
C;Genetics:
A;Gene: yviE
C;Superfamily: Bacillus subtilis hypothetical protein yviE

  Query Match             11.7%;  Score 78.5;  DB 2;  Length 191;
  Best Local Similarity   20.9%;  Pred. No. 3.3;
  Matches   31;  Conservative   25;  Mismatches   41;  Indels   51;  Gaps    7;

Qy          7 PPNICLDRLENWITSLAESQLQTRQQIKKLEELQQ---------------------KVSY 45
              |  : :|: : |      :|  :   |::||  |                      |:
Db         48 PGKLTIDQTQAW------EELDRKHVFKRIEEAAQQGHEDVMEGIARTAEEGDELMKIEN 101

Qy         46 KGDPIV-QHRPMLEERIVELFRNLMKS--------------AFVVERQPCMPMHPDRPLV 90
              ||:||  | |   |   ::|  |   |                :  |:| :   | :|:
Db        102 KGNPIASQARRNSEMHQIQLGENYAPSLSRVKIQYTPSQLDVQITPRKPVIQAEPHKPI- 160

Qy         91 IKTGVQFT---TKVRLLVKFPELNYQLK 115
                |::|    || :| ::|:||  ::
Db        161 ----VEYTPGNVKVDML-QYPDLNIDVE 183


Search completed: May 25, 2005, 17:45:25
Job time : 14.1698 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:15:30 ; Search time 53.6753 Seconds
(without alignments)
1230.701 Million cell updates/sec

Title: US-10-090-185-24
Perfect score: 669
Sequence: 1 IACIGGPPNICLDRLENWIT.....LNYQLKIKVCIDKDSGDVAA 129

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	669	100.0	163	2	Q9N145	Q9n145 macaca mula
2	669	100.0	770	1	STA3_BOVIN	P61635 bos taurus
3	669	100.0	770	1	STA3_HUMAN	P40763 homo sapien
4	669	100.0	770	1	STA3_MOUSE	P42227 mus musculu
5	666	99.6	770	1	STA3_RAT	P52631 rattus norv
6	665	99.4	771	2	Q6DV79	Q6dv79 gallus gall
7	658	98.4	769	2	Q9PVX8	Q9pvx8 xenopus lae
8	650	97.2	766	2	Q7ZXK3	Q7zxk3 xenopus lae
9	637	95.2	414	2	Q7ZTS5	Q7zts5 brachydanio
10	637	95.2	786	2	Q6NV46	Q6nv46 brachydanio
11	637	95.2	806	2	O93599	O93599 brachydanio
12	635	94.9	765	2	Q6DVF3	Q6dvf3 oryzias lat
13	635	94.9	785	2	Q6GUE7	Q6gue7 oryzias lat
14	624	93.3	764	2	Q90Y16	Q90y16 tetraodon f
15	600	89.7	767	2	O13133	O13133 oncorhynchu
16	555	83.0	161	2	Q704W5	Q704w5 bos taurus
17	446	66.7	751	2	Q8JGN0	Q8jgn0 xenopus lae
18	436	65.2	750	1	STA1_HUMAN	P42224 homo sapien
19	436	65.2	750	2	Q68D00	Q68d00 homo sapien
20	434	64.9	758	2	Q90Y17	Q90y17 tetraodon f
21	432	64.6	712	2	Q6P6Q7	Q6p6q7 rattus norv
22	432	64.6	749	2	Q9QXK0	Q9qxk0 rattus norv
23	430	64.3	757	2	Q764M5	Q764m5 sus scrofa
24	428	64.0	749	2	Q8C497	Q8c497 mus musculu
25	427	63.8	712	2	Q99K94	Q99k94 mus musculu
26	427	63.8	749	2	Q8C3V4	Q8c3v4 mus musculu
27	427	63.8	749	2	Q9D323	Q9d323 mus musculu
28	427	63.8	755	2	Q8C8M3	Q8c8m3 mus musculu
29	422	63.1	754	2	O13131	O13131 oncorhynchu
30	415	62.0	1153	2	Q8JFS5	Q8jfs5 brachydanio
31	410	61.3	754	2	O13132	O13132 oncorhynchu
32	409	61.1	528	2	Q8JFU8	Q8jfu8 brachydanio
33	406	60.7	718	2	Q801Y2	Q801y2 carassius a
34	395	59.0	749	1	STA1_MOUSE	P42225 mus musculu
35	386	57.7	749	2	O93598	O93598 brachydanio
36	386	57.7	749	2	Q6P943	Q6p943 brachydanio
37	379	56.7	748	1	STA4_HUMAN	Q14765 homo sapien
38	375	56.1	749	1	STA4_MOUSE	P42228 mus musculu
39	374	55.9	748	2	Q66HB2	Q66hb2 rattus norv
40	373	55.8	553	2	Q8JFU7	Q8jfu7 brachydanio
41	373	55.8	651	2	Q7ZZ53	Q7zz53 brachydanio
42	373	55.8	652	2	Q7ZZ77	Q7zz77 brachydanio
43	373	55.8	657	2	Q8AW24	Q8aw24 brachydanio
44	373	55.8	667	2	Q8AW20	Q8aw20 brachydanio
45	359	53.7	737	2	Q90Y15	Q90y15 tetraodon f

ALIGNMENTS

RESULT 1
Q9N145
ID Q9N145 PRELIMINARY; PRT; 163 AA.
AC Q9N145;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Signal transducer and activator of transcription (Fragment).
GN Name=STAT3;
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Arredondo J.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF227560; AAF73401.1; -.
DR HSSP; P42227; 1BG1.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR008967; P53_like_DNA_bnd.
DR InterPro; IPR001217; STAT.
DR Pfam; PF01017; STAT_alpha; 1.
DR Pfam; PF02864; STAT_bind; 1.
FT NON_TER 1 1
FT NON_TER 163 163
SQ SEQUENCE 163 AA; 18855 MW; DD3F11E7F5B9878C CRC64;

Query Match 100.0%; Score 669; DB 2; Length 163;
Best Local Similarity 100.0%; Pred. No. 7.6e-55;
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         19 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 78

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         79 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 138

Qy        121 DKDSGDVAA 129
              |||||||||
Db        139 DKDSGDVAA 147
```

RESULT 2
STA3_BOVIN
ID STA3_BOVIN STANDARD; PRT; 770 AA.
AC P61635;

```
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Signal transducer and activator of transcription 3.
GN   Name=STAT3;
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Mammary gland;
RA   Seyfert H.M., Wheeler T.T., Moolenaar A., Pitra C.;
RT   "The STAT5B-encoding gene was flipped across the STAT3/STAT5A-locus
RT   during ruminant evolution.";
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes (By similarity).
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation (By similarity).
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ620655; CAF06182.1; -.
DR   PROSITE; PS50001; SH2; 1.
KW   Activator; DNA-binding; Nuclear protein; Phosphorylation; SH2 domain;
KW   Transcription regulation.
FT   DOMAIN      580    670       SH2.
FT   MOD_RES     705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES     727    727       Phosphoserine (By similarity).
SQ   SEQUENCE   770 AA;  87974 MW;  9CEB147C73E83274 CRC64;

  Query Match            100.0%;  Score 669;  DB 1;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 4.6e-54;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


RESULT 3
STA3_HUMAN
ID   STA3_HUMAN     STANDARD;      PRT;   770 AA.
AC   P40763; O14916; Q9BW54;
DT   01-FEB-1995 (Rel. 31, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Signal transducer and activator of transcription 3 (Acute-phase
DE   response factor).
GN   Name=STAT3; Synonyms=APRF;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Placenta;
RX   MEDLINE=94208062; PubMed=7512451; DOI=10.1016/0092-8674(94)90235-6;
RA   Akira S., Nishio Y., Inoue M., Wang X.-J., Wei S., Matsusaka T.,
RA   Yoshida K., Sudo T., Naruto M., Kishimoto T.;
RT   "Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-
RT   related transcription factor involved in the gp130-mediated signaling
RT   pathway.";
RL   Cell 77:63-71(1994).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=98296260; PubMed=9630560; DOI=10.1016/S0378-1119(98)00185-1;
RA   Della Pietra L., Bressan A., Pezzotti A., Serlupi-Crescenzi O.;
RT   "Highly conserved amino-acid sequence between murine STAT3 and a
RT   revised human STAT3 sequence.";
RL   Gene 213:119-124(1998).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM DEL-701), AND VARIANT ILE-143.
RA   Rieder M.J., Daniels R.L., da Ponte S.H., Hastings N.C., Ahearn M.O.,
RA   Rajkumar N., Yi Q., Nickerson D.A.;
RT   "SeattleSNPs. NHLBI HL66682 program for genomic applications, UW-
RT   FHCRC, Seattle, WA (URL: http://pga.gs.washington.edu).";
RL   Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND DEL-701).
RC   TISSUE=Kidney, and Pancreas;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   SEQUENCE OF 564-704 FROM N.A.
RC   TISSUE=Liver;
RA   Della Pietra L., Bressan A., Pezzotti A.R., Serlupi-Crescenzi O.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   PHOSPHORYLATION ON SERINE.
RX   MEDLINE=95215843; PubMed=7701321;
RA   Zhang X., Blenis J., Li H.-C., Schindler C., Chen-Kiang S.;
RT   "Requirement of serine phosphorylation for formation of STAT-promoter
RT   complexes.";
RL   Science 267:1990-1994(1995).
RN   [7]
RP   INTERACTION WITH NCOA1.
RX   PubMed=11773079; DOI=10.1074/jbc.M111486200;
RA   Giraud S., Bienvenu F., Avril S., Gascan H., Heery D.M., Coqueret O.;
RT   "Functional interaction of STAT3 transcription factor with the
RT   coactivator NcoA/SRC1a.";
RL   J. Biol. Chem. 277:8004-8011(2002).
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes.
```

```
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P40763-1; Sequence=Displayed;
CC       Name=Del-701;
CC         IsoId=P40763-2; Sequence=VSP_010474;
CC   -!- TISSUE SPECIFICITY: Heart, brain, placenta, lung, liver, skeletal
CC       muscle, kidney and pancreas.
CC   -!- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,
CC       LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation
CC       is important for the formation of stable DNA-binding STAT3
CC       homodimers and maximal transcriptional activity.
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L29277; AAA58374.1; -.
DR   EMBL; AJ012463; CAA10032.1; -.
DR   EMBL; AY572796; AAS66986.1; -.
DR   EMBL; BC000627; AAH00627.1; -.
DR   EMBL; BC014482; AAH14482.1; -.
DR   EMBL; AF029311; AAB84254.1; -.
DR   PIR; A54444; A54444.
DR   HSSP; P42227; 1BG1.
DR   TRANSFAC; T01493; -.
DR   Genew; HGNC:11364; STAT3.
DR   H-InvDB; HIX0013840; -.
DR   MIM; 102582; -.
DR   GO; GO:0005737; C:cytoplasm; TAS.
DR   GO; GO:0005634; C:nucleus; TAS.
DR   GO; GO:0005062; F:hematopoietin/interferon-class (D200-domain. . .; TAS.
DR   GO; GO:0003700; F:transcription factor activity; TAS.
DR   GO; GO:0006928; P:cell motility; TAS.
DR   GO; GO:0007259; P:JAK-STAT cascade; TAS.
DR   GO; GO:0000122; P:negative regulation of transcription from P. . .; TAS.
DR   GO; GO:0007399; P:neurogenesis; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
KW   Activator; Alternative splicing; DNA-binding; Nuclear protein;
KW   Phosphorylation; Polymorphism; SH2 domain; Transcription regulation.
FT   DOMAIN      580    670       SH2.
FT   MOD_RES     705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES     727    727       Phosphoserine (By similarity).
FT   VARSPLIC    701    701       Missing (in isoform Del-701).
FT                                /FTId=VSP_010474.
FT   VARIANT      32     32       Q -> K (in dbSNP:1803125).
FT                                /FTId=VAR_018683.
FT   VARIANT     143    143       M -> I.
FT                                /FTId=VAR_018679.
FT   CONFLICT    288    288       Q -> H (in Ref. 1).
FT   CONFLICT    460    460       P -> S (in Ref. 1).
FT   CONFLICT    548    548       K -> N (in Ref. 1).
FT   CONFLICT    561    561       F -> Y (in Ref. 1).
FT   CONFLICT    667    667       V -> L (in Ref. 1).
FT   CONFLICT    730    730       T -> A (in Ref. 1).
SQ   SEQUENCE   770 AA;  88067 MW;  6C00632211C8012D CRC64;

  Query Match             100.0%;  Score 669;  DB 1;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-54;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


RESULT 4
STA3_MOUSE
ID   STA3_MOUSE     STANDARD;      PRT;   770 AA.
AC   P42227;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Signal transducer and activator of transcription 3 (Acute-phase
DE   response factor).
GN   Name=Stat3; Synonyms=Aprf;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A), AND SEQUENCE OF 154-158; 181-185
RP   AND 632-640.
RC   STRAIN=BALB/c; TISSUE=Liver;
RX   MEDLINE=94208062; PubMed=7512451; DOI=10.1016/0092-8674(94)90235-6;
RA   Akira S., Nishio Y., Inoue M., Wang X.J., Shi W., Matsusaka T.,
RA   Yoshida K., Sudo T., Naruto M., Kishimoto T.;
RT   "Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-
RT   related transcription factor involved in the gp130-mediated signaling
RT   pathway.";
RL   Cell 77:63-71(1994).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   TISSUE=Thymus;
RX   MEDLINE=94188718; PubMed=8140422;
RA   Zhong Z., Wen Z., Darnell J.E. Jr.;
RT   "Stat3: a STAT family member activated by tyrosine phosphorylation in
RT   response to epidermal growth factor and interleukin-6.";
RL   Science 264:95-98(1994).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   TISSUE=Brain;
RX   MEDLINE=95014185; PubMed=7523373;
RA   Raz R., Durbin J.E., Levy D.E.;
RT   "Acute phase response factor and additional members of the interferon-
RT   stimulated gene factor 3 family integrate diverse signals from
RT   cytokines, interferons, and growth factors.";
RL   J. Biol. Chem. 269:24391-24395(1994).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3B).
RC   STRAIN=BALB/c,. and C57BL/6; TISSUE=Liver;
RX   MEDLINE=96016116; PubMed=7568080;
RA   Schaefer T.S., Sanders L.K., Nathans D.;
RT   "Cooperative transcriptional activity of Jun and Stat3 beta, a short
RT   form of Stat3.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:9097-9101(1995).
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   STRAIN=129/SvJ;
```

```
RX   PubMed=11161808; DOI=10.1006/geno.2000.6433;
RA   Miyoshi K., Cui Y., Riedlinger G., Lehoczky J., Zon L., Oka T.,
RA   Dewar K., Hennighausen L.;
RT   "Structure of the mouse stat 3/5 locus: evolution from Drosophila to
RT   zebrafish to mouse.";
RL   Genomics 71:150-155(2001).
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   STRAIN=C57BL/6J, and NOD/LtJ;
RA   Davoodi-Semiromi A., She J.-X.;
RT   "A mutant Stat5b with weaker DNA binding defines a key defective
RT   pathway in non-obese diabetic (NOD) mice.";
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   STRAIN=FVB/N; TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [8]
RP   PHOSPHORYLATION SITE SER-727, AND MUTAGENESIS.
RX   MEDLINE=95354205; PubMed=7543024; DOI=10.1016/0092-8674(95)90311-9;
RA   Wen Z., Zhong Z., Darnell J.E. Jr.;
RT   "Maximal activation of transcription by Stat1 and Stat3 requires both
RT   tyrosine and serine phosphorylation.";
RL   Cell 82:241-250(1995).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 136-716.
RX   MEDLINE=98334373; PubMed=9671298; DOI=10.1038/28101;
RA   Becker S., Groner B., Mueller C.W.;
RT   "Three-dimensional structure of the Stat3beta homodimer bound to
RT   DNA.";
RL   Nature 394:145-151(1998).
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes. STAT3B interacts with the N-terminal
CC       part of JUN to activate such promoters in a cooperative way.
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=Stat3A;
CC         IsoId=P42227-1; Sequence=Displayed;
CC       Name=Stat3B;
CC         IsoId=P42227-2; Sequence=VSP_006287;
CC       Name=Del-701;
CC         IsoId=P42227-3; Sequence=VSP_010475;
CC   -!- TISSUE SPECIFICITY: STAT3A is seen in the liver, spleen, and
CC       kidney. STAT3B is also detected in the liver, although in a much
CC       less abundant manner.
CC   -!- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,
CC       LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation
CC       is important for the formation of stable DNA-binding STAT3
CC       homodimers and maximal transcriptional activity (By similarity).
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L29278; AAA37254.1; -.
DR   EMBL; U06922; AAA19452.1; -.
DR   EMBL; U08378; AAA56668.1; -.
DR   EMBL; U30709; AAC52612.1; -.
DR   EMBL; AF246978; AAL59017.1; -.
DR   EMBL; AY299489; AAQ75418.1; -.
DR   EMBL; AY299490; AAQ75419.1; -.
DR   EMBL; BC003806; AAH03806.1; -.
DR   PIR; I49508; I49508.
DR   PDB; 1BG1; X-ray; A=1-722.
DR   TRANSFAC; T01574; -.
DR   MGD; MGI:103038; Stat3.
DR   GO; GO:0005737; C:cytoplasm; IDA.
DR   GO; GO:0005634; C:nucleus; IDA.
DR   GO; GO:0005886; C:plasma membrane; IDA.
DR   GO; GO:0003677; F:DNA binding; IDA.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   GO; GO:0016563; F:transcriptional activator activity; IDA.
DR   GO; GO:0007259; P:JAK-STAT cascade; IDA.
DR   GO; GO:0006357; P:regulation of transcription from Pol II pro. . .; IDA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
KW   3D-structure; Activator; Acute phase; Alternative splicing;
KW   Direct protein sequencing; DNA-binding; Nuclear protein;
KW   Phosphorylation; SH2 domain; Transcription regulation.
FT   DOMAIN       580    670       SH2.
FT   MOD_RES      705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES      727    727       Phosphoserine.
FT   VARSPLIC     716    770       TTCSNTIDLPMSPRTLDSLMQFGNNGEGAEPSAGGQFESLT
FT                                 FDMDLTSECATSPM -> FIDAVWK (in isoform
FT                                 Stat3B).
FT                                 /FTId=VSP_006287.
FT   VARSPLIC     701    701       Missing (in isoform Del-701).
FT                                 /FTId=VSP_010475.
FT   MUTAGEN      727    727       S->A: Decreased transcriptional
FT                                 activation.
FT   CONFLICT      16     16       E -> K (in Ref. 2).
FT   CONFLICT      25     25       S -> T (in Ref. 2 and 4).
FT   CONFLICT     394    394       M -> I (in Ref. 1).
FT   HELIX        139    180
FT   TURN         181    182
FT   TURN         197    198
FT   HELIX        199    237
FT   TURN         238    238
FT   HELIX        239    251
FT   TURN         252    253
FT   HELIX        261    290
FT   TURN         294    295
FT   TURN         297    301
FT   HELIX        302    320
FT   STRAND       321    328
FT   TURN         330    331
FT   TURN         333    334
FT   TURN         336    337
```

```
FT   STRAND      338    340
FT   TURN        341    342
FT   STRAND      345    351

  Query Match             100.0%;  Score 669;  DB 1;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-54;
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


RESULT 5
STA3_RAT
ID   STA3_RAT       STANDARD;      PRT;   770 AA.
AC   P52631;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Signal transducer and activator of transcription 3.
GN   Name=Stat3;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=96102059; PubMed=8530402; DOI=10.1074/jbc.270.50.29998;
RA   Ripperger J.A., Fritz S., Richter K., Hocke G.M., Lottspeich F.,
RA   Fey G.H.;
RT   "Transcription factors Stat3 and Stat5b are present in rat liver
RT   nuclei late in an acute phase response and bind interleukin-6 response
RT   elements.";
RL   J. Biol. Chem. 270:29998-30006(1995).
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes.
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation (By similarity).
CC   -!- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,
CC       LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation
CC       is important for the formation of stable DNA-binding STAT3
CC       homodimers and maximal transcriptional activity (By similarity).
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X91810; CAA62920.1; -.
DR   HSSP; P42227; 1BG1.
DR   RGD; 3772; Stat3.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
KW   Activator; DNA-binding; Nuclear protein; Phosphorylation; SH2 domain;
KW   Transcription regulation.
FT   DOMAIN      580    670       SH2.
FT   MOD_RES     705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES     727    727       Phosphoserine (By similarity).
SQ   SEQUENCE   770 AA;  88039 MW;  D74A0C76954754ED CRC64;

  Query Match             99.6%;  Score 666;  DB 1;  Length 770;
  Best Local Similarity   99.2%;  Pred. No. 8.7e-54;
  Matches  128;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


RESULT 6
Q6DV79
ID   Q6DV79      PRELIMINARY;      PRT;   771 AA.
AC   Q6DV79;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Signal transducer and activator of transcription 3.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zhou G.Y., Leung F.C.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY641397; AAT64887.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   771 AA;  88174 MW;  71AC855C5DEC03E2 CRC64;

  Query Match             99.4%;  Score 665;  DB 2;  Length 771;
  Best Local Similarity   99.2%;  Pred. No. 1.1e-53;
  Matches  128;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
```

```
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVFKTGVQFTTKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


RESULT 7
Q9PVX8
ID   Q9PVX8      PRELIMINARY;      PRT;   769 AA.
AC   Q9PVX8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat 3.
GN   Name=stat 3;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20107399; PubMed=10642787; DOI=10.1006/dbio.1999.9518;
RA   Nishinakamura R., Matsumoto Y., Matsuda T., Ariizumi T., Heike T.,
RA   Asashima M., Yokota T.;
RT   "Activation of Stat3 by cytokine receptor gp130 ventralizes Xenopus
RT   embryos independent of BMP-4.";
RL   Dev. Biol. 216:481-490(1999).
DR   EMBL; AB017701; BAA86061.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   769 AA;  87974 MW;  0905C03263303069 CRC64;

  Query Match             98.4%;  Score 658;  DB 2;  Length 769;
  Best Local Similarity   98.4%;  Pred. No. 4.9e-53;
  Matches  127;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQIRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTNKVRLLVKFPELNYQLKIKVCI 368

Qy        121 DKDSGDVAA 129
              |||||||||
Db        369 DKDSGDVAA 377


RESULT 8
Q7ZXK3
ID   Q7ZXK3      PRELIMINARY;      PRT;   766 AA.
AC   Q7ZXK3;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat3-A protein.
GN   Name=stat3-A;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC044717; AAH44717.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   766 AA;  87599 MW;  31018A3321CCEB9C CRC64;

  Query Match             97.2%;  Score 650;  DB 2;  Length 766;
  Best Local Similarity   96.9%;  Pred. No. 2.7e-52;
  Matches  125;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              |||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        249 IACIGGPPNICLDRLENWITSLAESQLQTRQQIRKLEELQQKVSYKGDPIVQHRPMLEER 308

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        309 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTNKVRLLVKFPELNYQLKIKVCI 368
```

```
Qy        121 DKDSGDVAA 129
              |||||: ||
Db        369 DKDSGEGAA 377


RESULT 9
Q7ZTS5
ID   Q7ZTS5      PRELIMINARY;      PRT;   414 AA.
AC   Q7ZTS5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat3 protein.
GN   Name=stat3;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB; TISSUE=Whole body;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB; TISSUE=Whole body;
RA   Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC045276; AAH45276.1; -.
DR   HSSP; P42227; 1BG1.
DR   ZFIN; ZDB-GENE-980526-68; stat3.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
SQ   SEQUENCE   414 AA;  48253 MW;  0FFD1B509B7526BD CRC64;

  Query Match             95.2%;  Score 637;  DB 2;  Length 414;
  Best Local Similarity   93.8%;  Pred. No. 2.2e-51;
  Matches  121;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              |||||||||||||||| |||||||||| ||||:||||||||||||||:|||| |||:
Db        250 IACIGGPPNICLDRLETWITSLAESQLQIRQQIRKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 369

Qy        121 DKDSGDVAA 129
              ||:||||||
Db        370 DKESGDVAA 378


RESULT 10
Q6NV46
ID   Q6NV46      PRELIMINARY;      PRT;   786 AA.
AC   Q6NV46;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Stat3 protein.
GN   Name=stat3;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC068320; AAH68320.1; -.
DR   ZFIN; ZDB-GENE-980526-68; stat3.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   786 AA;  90039 MW;  FC7371D0B0E5447E CRC64;

  Query Match             95.2%;  Score 637;  DB 2;  Length 786;
  Best Local Similarity   93.8%;  Pred. No. 4.7e-51;
  Matches  121;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              |||||||||||||||| |||||||||| ||||:||||||||||||||:|||| |||:
```

```
Db        250 IACIGGPPNICLDRLETWITSLAESQLQIRQQIRKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 369

Qy        121 DKDSGDVAA 129
              ||:||||||
Db        370 DKESGDVAA 378


RESULT 11
O93599
ID   O93599      PRELIMINARY;      PRT;   806 AA.
AC   O93599;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Transcription factor.
GN   Name=stat3;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Oates A.C.;
RL   Thesis (1998), University of Melbourne, Australia.
DR   EMBL; AJ005693; CAA06677.1; -.
DR   HSSP; P42227; 1BG1.
DR   ZFIN; ZDB-GENE-980526-68; stat3.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   806 AA;  92151 MW;  74BC4EA401C3C942 CRC64;

  Query Match             95.2%;  Score 637;  DB 2;  Length 806;
  Best Local Similarity   93.8%;  Pred. No. 4.8e-51;
  Matches  121;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              |||||||||||||||| |||||||||| ||||:|||||||||||||||:|||| |||:
Db        250 IACIGGPPNICLDRLETWITSLAESQLQIRQQIRKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 369

Qy        121 DKDSGDVAA 129
              ||:||||||
Db        370 DKESGDVAA 378


RESULT 12
Q6DVF3
ID   Q6DVF3      PRELIMINARY;      PRT;   765 AA.
AC   Q6DVF3;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Signal transducer and activation of transcription factor 3.
OS   Oryzias latipes (Medaka fish) (Japanese ricefish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Liu R., Hong Y.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY639947; AAT64912.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   765 AA;  87566 MW;  F5D01408748EC703 CRC64;

  Query Match             94.9%;  Score 635;  DB 2;  Length 765;
  Best Local Similarity   93.8%;  Pred. No. 7e-51;
  Matches  121;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              |||||||||||||||| |||||||||| |||||||||||||||||||||:|||| |||:
Db        250 IACIGGPPNICLDRLETWITSLAESQLQIRQQIKKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||:||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        310 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTNKVRLLVKFPELNYQLKIKVCI 369

Qy        121 DKDSGDVAA 129
              ||:||||||
Db        370 DKESGDVAA 378


RESULT 13
Q6GUE7
ID   Q6GUE7      PRELIMINARY;      PRT;   785 AA.
AC   Q6GUE7;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Signal transducer and activator of transcription 3 isoform 1.
OS   Oryzias latipes (Medaka fish) (Japanese ricefish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Liu R., Hong Y.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY641434; AAT46364.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
```

```
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   785 AA;  89643 MW;  81F231BDE27DE938 CRC64;

  Query Match             94.9%;  Score 635;  DB 2;  Length 785;
  Best Local Similarity   93.8%;  Pred. No. 7.2e-51;
  Matches  121;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              |||||||||||||||| |||||||||| ||||||||||||||||||||:|||| |||:
Db        250 IACIGGPPNICLDRLETWITSLAESQLQIRQQIKKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||:|||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        310 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTNKVRLLVKFPELNYQLKIKVCI 369

Qy        121 DKDSGDVAA 129
              ||:||||||
Db        370 DKESGDVAA 378


RESULT 14
Q90Y16
ID   Q90Y16      PRELIMINARY;      PRT;   764 AA.
AC   Q90Y16;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   STAT3.
OS   Tetraodon fluviatilis (Puffer fish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=47145;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sung S.C., Fan T.J., Leu J.H., Chou C.M., Huang C.J.;
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF307106; AAL09415.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   764 AA;  87411 MW;  E661FFE18BEFD8BE CRC64;

  Query Match             93.3%;  Score 624;  DB 2;  Length 764;
  Best Local Similarity   92.2%;  Pred. No. 7.5e-50;
  Matches  119;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              |||||||||||||||| |||||||||| ||||||||||||||||||||:|||| |||:
Db        250 IACIGGPPNICLDRLETWITSLAESQLQIRQQIKKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||:|||||||||||||||||||||||||||||||| |||||||| ||||||||||||
Db        310 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTNKVRLLVKFQELNYQLKIKVCI 369

Qy        121 DKDSGDVAA 129
              ||:|||||:
Db        370 DKESGDVAS 378


RESULT 15
O13133
ID   O13133      PRELIMINARY;      PRT;   767 AA.
AC   O13133;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat3.
GN   Name=rbtStat3;
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8022;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Johnson M.C., Mourich D.V., Leong J.C.;
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U60333; AAB60926.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
SQ   SEQUENCE   767 AA;  87816 MW;  FED97740C74C3798 CRC64;

  Query Match             89.7%;  Score 600;  DB 2;  Length 767;
  Best Local Similarity   89.9%;  Pred. No. 1.3e-47;
  Matches  116;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 60
              |||||||| ||||||| ||||| | ||| ||||||||||||||||||||:|||| |||:
Db        250 IACIGGPPKICLDRLETWITSLGEIQLQIRQQIKKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy         61 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPELNYQLKIKVCI 120
              ||:|||||||||||| |||||||||||||||||| |||||||||||||||||||| |
Db        310 IVDLFRNLMKSAFVVERQRCMPMHPDRPLVIKTGVQFTNKVRLLVKFPELNYQLKIKVII 369

Qy        121 DKDSGDVAA 129
              ||:||||||
Db        370 DKESGDVAA 378


Search completed: May 25, 2005, 17:43:41
Job time : 53.6753 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:14:45 ; Search time 51.0983 Seconds
(without alignments)
726.619 Million cell updates/sec

Title: US-10-090-185-25
Perfect score: 494
Sequence: 1 KKLEELQQKVSYKGDPIVQH.....LNYQLKIKVCIDKDSGDVAA 96

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	494	100.0	96	4	AAY72857	Aay72857 Mouse Sta
2	494	100.0	129	4	AAY72856	Aay72856 Mouse Sta
3	494	100.0	185	4	AAY72855	Aay72855 Mouse Sta
4	494	100.0	223	4	AAY72854	Aay72854 Mouse Sta
5	494	100.0	271	4	AAY72841	Aay72841 Mouse Sta
6	494	100.0	720	5	AAE22055	Aae22055 Human Sta
7	494	100.0	769	5	ABB57164	Abb57164 Mouse isc
8	494	100.0	769	5	AAE22054	Aae22054 Human Sta
9	494	100.0	769	5	AAE22056	Aae22056 Human pro
10	494	100.0	770	2	AAR72082	Aar72082 Mouse Sta
11	494	100.0	770	2	AAR82995	Aar82995 Mouse liv
12	494	100.0	770	2	AAW03176	Aaw03176 Mouse STA
13	494	100.0	770	2	AAY03768	Aay03768 Human STA
14	494	100.0	770	3	AAB12377	Aab12377 N-termina
15	494	100.0	770	5	AAE14652	Aae14652 Murine ST
16	494	100.0	770	5	ABG69497	Abg69497 Human bai
17	494	100.0	770	6	ABU10476	Abu10476 Mouse STA
18	494	100.0	770	8	ADN04365	Adn04365 Antipsori
19	494	100.0	770	8	ADP54789	Adp54789 Human PRO
20	494	100.0	793	3	AAB58442	Aab58442 Lung canc
21	491	99.4	770	7	ADD44738	Add44738 Rat Prote
22	489	99.0	770	2	AAR82993	Aar82993 Human pla
23	489	99.0	770	4	AAB19964	Aab19964 Human sig
24	489	99.0	770	5	AAE15174	Aae15174 Human Sta
25	489	99.0	770	7	ADD44740	Add44740 Human Pro
26	396	80.2	228	4	AAY72861	Aay72861 Mouse Sta
27	396	80.2	229	4	AAY72850	Aay72850 Mouse Sta
28	396	80.2	229	4	AAY72863	Aay72863 Mouse Sta
29	396	80.2	229	4	AAY72862	Aay72862 Mouse Sta
30	396	80.2	252	4	AAY72846	Aay72846 Mouse Sta
31	379	76.7	229	4	AAY72860	Aay72860 Mouse Sta
32	318	64.4	213	4	AAY72851	Aay72851 Mouse Sta
33	318	64.4	236	4	AAY72847	Aay72847 Mouse Sta
34	311	63.0	749	5	AAG78526	Aag78526 Rat STAT-
35	310	62.8	268	4	AAY72844	Aay72844 Mouse Sta
36	310	62.8	582	2	AAW62996	Aaw62996 Human tru
37	310	62.8	582	6	ABU04748	Abu04748 Human exp
38	310	62.8	582	8	ADH57036	Adh57036 Truncated
39	310	62.8	680	6	ABR59713	Abr59713 Human sig
40	310	62.8	701	2	AAR41335	Aar41335 84 kD ISG
41	310	62.8	701	6	ABU04741	Abu04741 Human exp
42	310	62.8	712	2	AAR72079	Aar72079 Human Sta
43	310	62.8	712	2	AAW03170	Aaw03170 Human STA
44	310	62.8	712	2	AAW62995	Aaw62995 Human Sta
45	310	62.8	712	6	ABU04747	Abu04747 Human exp

ALIGNMENTS

RESULT 1
AAY72857
ID AAY72857 standard; protein; 96 AA.
XX
AC AAY72857;
XX
DT 31-MAY-2001 (first entry)
XX
DE Mouse Stat3 protein fragment #15 (282-377 amino acids).
XX
KW Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW cellular transformation; dysproliferative disease; cancer; psoriasis;
KW therapy.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT Region 61. .77
FT /note= "Stat3-c-Jun interaction region 2; corresponds to
FT 342-358 position of Stat3 protein"
XX
PN WO200116605-A2.
XX
PD 08-MAR-2001.
XX
PF 30-AUG-2000; 2000WO-US023822.
XX
PR 31-AUG-1999; 99US-00387418.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI Zhang X, Horvath C, Wrzeszcynska MH, Darnell JE;
XX
DR WPI; 2001-226705/23.
XX
PT Identifying an agent for use in modulating the interaction between
PT transcription factor c-Jun and a Stat3 protein.
XX
PS Claim 65; Page 81; 86pp; English.
XX
CC The present sequence is mouse Stat3 protein fragment containing 282-377
CC amino acids of Stat3 protein. The invention relates to methods for
CC identifying interacting regions of transcription factors and methods for
CC identifying agents which modulates the interaction between a
CC transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC Stat-3, useful for modulating gene transcription e.g., cellular
CC transformation. These identifying agents are used in the treatment of
CC dysproliferative diseases and also for treating cancer and psoriasis. A

```
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 96 AA;

  Query Match           100.0%;  Score 494;  DB 4;  Length 96;
  Best Local Similarity  100.0%;  Pred. No. 3.6e-55;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96


RESULT 2
AAY72856
ID   AAY72856 standard; protein; 129 AA.
XX
AC   AAY72856;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 protein fragment #14 (249-377 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          94. .110
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 80-81; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 249-377
CC   amino acids of Stat3 protein. The invention relates to methods for
CC   identifying interacting regions of transcription factors and methods for
CC   identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 129 AA;

  Query Match           100.0%;  Score 494;  DB 4;  Length 129;
  Best Local Similarity  100.0%;  Pred. No. 5.2e-55;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         34 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 93

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db         94 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 129


RESULT 3
AAY72855
ID   AAY72855 standard; protein; 185 AA.
XX
AC   AAY72855;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 protein fragment #13 (193-377 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          150. .166
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 80; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 193-377
CC   amino acids of Stat3 protein. The invention relates to methods for
CC   identifying interacting regions of transcription factors and methods for
CC   identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 185 AA;

  Query Match           100.0%;  Score 494;  DB 4;  Length 185;
  Best Local Similarity  100.0%;  Pred. No. 8.4e-55;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 149

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
```

```
                    ||||||||||||||||||||||||||||||||||||
Db        150 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 185


RESULT 4
AAY72854
ID   AAY72854 standard; protein; 223 AA.
XX
AC   AAY72854;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 protein fragment #12 (155-377 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          188. .204
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 79; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 155-377
CC   amino acids of Stat3 protein. The invention relates to methods for
CC   identifying interacting regions of transcription factors and methods for
CC   identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 223 AA;

  Query Match             100.0%;  Score 494;  DB 4;  Length 223;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-54;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 187

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        188 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 223


RESULT 5
AAY72841
ID   AAY72841 standard; protein; 271 AA.
XX
AC   AAY72841;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 protein fragment #2 (107-377 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          24. .48
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
FT   Region          236. .252
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 67-68; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 107-377
CC   amino acids of Stat3 protein. This Stat3 fragment showed strong binding
CC   to c-Jun protein in the cell extract. The invention relates to methods
CC   for identifying interacting regions of transcription factors and methods
CC   for identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 271 AA;

  Query Match             100.0%;  Score 494;  DB 4;  Length 271;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-54;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 235

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        236 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271


RESULT 6
AAE22055
ID   AAE22055 standard; protein; 720 AA.
XX
AC   AAE22055;
XX
```

DT 25-JUL-2002 (first entry)
XX
DE Human Stat3beta protein.
XX
KW Human; signal transducer and activator of transcription 3; ischaemia;
KW immune response; Stat3; coronary atherosclerosis; vascular occlusion;
KW hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;
KW inflammation; chronic obstructive pulmonary disease; cardiac arrest;
KW insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;
KW shock; chronic active hepatitis; adult respiratory distress syndrome;
KW nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;
KW Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;
KW polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;
KW proliferative disease; Grave's disease; ulcerative colitis; sarcoma;
KW carcinoma; degenerative disorder; gene therapy; growth deficiency;
KW cirrhosis; hypoproliferative disorder; lesion; Statbeta.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 713. .714
FT /note= "Encoded by ACA CCA TTC"
XX
PN WO200220032-A1.
XX
PD 14-MAR-2002.
XX
PF 10-SEP-2001; 2001WO-US028254.
XX
PR 08-SEP-2000; 2000US-0231212P.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
PA (UYSF-) UNIV SOUTH FLORIDA.
XX
PI Yu H, Pardoll D, Jove R, Dalton W;
XX
DR WPI; 2002-362218/39.
DR N-PSDB; AAD35066.
XX
PT Modulating angiogenesis and an immune response in an individual, for
PT treating a hypoxic or ischemic condition, comprises administering a
PT compound that modulates the activity of a signal transducer and activator
PT of transcription 3.
XX
PS Disclosure; Page 87-89; 94pp; English.
XX
CC The invention relates to a method of modulating angiogenesis and immune
CC response. Method involves administering to an individual a compound that
CC modulate the activity of signal transducer and activator of transcription
CC 3 (Stat3). Modulating angiogenesis is useful for treating or preventing
CC hypoxic or ischaemic condition or disorder which is the result of stroke,
CC ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,
CC tissue ischaemia in the lower extremities, infarction, trauma, vascular
CC occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,
CC chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,
CC epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest,
CC nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy
CC with neovascularisation. Suppressing an immune response is useful for
CC ameliorating a symptom of an autoimmune disease such as systemic lupus
CC erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,
CC Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,
CC mixed connective tissue disease, primary biliary cirrhosis, pernicious
CC anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,
CC gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,
CC idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's
CC disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune
CC infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and
CC dense deposit disease. The method is useful in preventing or treating
CC specific proliferative and oncogenic disease which includes sarcomas and
CC carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,
CC fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,
CC hypoproliferative disorders, physical trauma, lesions and wounds. The
CC method is also used in gene therapy. The present sequence is human
CC Stat3beta protein
XX
SQ Sequence 720 AA;

Query Match 100.0%; Score 494; DB 5; Length 720;
Best Local Similarity 100.0%; Pred. No. 4.8e-54;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
Db 282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy 61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
Db 342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377

RESULT 7
ABB57164
ID ABB57164 standard; protein; 769 AA.
XX
AC ABB57164;
XX
DT 07-MAR-2002 (first entry)
XX
DE Mouse ischaemic condition related protein sequence SEQ ID NO:398.
XX
KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease.
XX
OS Mus musculus.
XX
PN WO200188188-A2.
XX
PD 22-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-JP004192.
XX
PR 18-MAY-2000; 2000JP-00145977.
XX
PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
DR WPI; 2002-034733/04.
DR N-PSDB; ABI99454.
XX
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or by
PT determining the expression profile of a gene group comprising these
PT genes.
XX
PS Claim 2; Page 1084-1087; 2690pp; English.
XX
CC The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (ABI99202 to ABI99912, encoding the
CC protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The expression
CC levels or expression profiles produced by these genes are used as an
CC indicator when screening for ischaemic condition-improving drugs or
CC therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR
CC primers for a mouse ischaemic condition related sequence, which are used
CC in the exemplification of the present invention
XX
SQ Sequence 769 AA;

Query Match 100.0%; Score 494; DB 5; Length 769;
Best Local Similarity 100.0%; Pred. No. 5.3e-54;

```
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              |||||||||||||||||||||||||||||||||||
Db    .    342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 8
AAE22054
ID   AAE22054 standard; protein; 769 AA.
XX
AC   AAE22054;
XX
DT   25-JUL-2002  (first entry)
XX
DE   Human Stat3 protein.
XX
KW   Human; signal transducer and activator of transcription 3; ischaemia;
KW   immune response; Stat3; coronary atherosclerosis; vascular occlusion;
KW   hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;
KW   inflammation; chronic obstructive pulmonary disease; cardiac arrest;
KW   insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;
KW   shock; chronic active hepatitis; adult respiratory distress syndrome;
KW   nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;
KW   Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;
KW   polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;
KW   proliferative disease; Grave's disease; ulcerative colitis; sarcoma;
KW   carcinoma; degenerative disorder; gene therapy; growth deficiency;
KW   cirrhosis; hypoproliferative disorder; lesion.
XX
OS   Homo sapiens.
XX
PN   WO200220032-A1.
XX
PD   14-MAR-2002.
XX
PF   10-SEP-2001; 2001WO-US028254.
XX
PR   08-SEP-2000; 2000US-0231212P.
XX
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (UYSF-) UNIV SOUTH FLORIDA.
XX
PI   Yu H,  Pardoll D,  Jove R,  Dalton W;
XX
DR   WPI; 2002-362218/39.
DR   N-PSDB; AAD35065.
XX
PT   Modulating angiogenesis and an immune response in an individual, for
PT   treating a hypoxic or ischemic condition, comprises administering a
PT   compound that modulates the activity of a signal transducer and activator
PT   of transcription 3.
XX
PS   Disclosure; Page 83-85; 94pp; English.
XX
CC   The invention relates to a method of modulating angiogenesis and immune
CC   response. Method involves administering to an individual a compound that
CC   modulate the activity of signal transducer and activator of transcription
CC   3 (Stat3). Modulating angiogenesis is useful for treating or preventing
CC   hypoxic or ischaemic condition or disorder which is the result of stroke,
CC   ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,
CC   tissue ischaemia in the lower extremities, infarction, trauma, vascular
CC   occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,
CC   chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,
CC   epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest,
CC   nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy
CC   with neovascularisation. Suppressing an immune response is useful for
CC   ameliorating a symptom of an autoimmune disease such as systemic lupus
CC   erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,
CC   Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,
CC   mixed connective tissue disease, primary biliary cirrhosis, pernicious
CC   anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,
CC   gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,
CC   idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's
CC   disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune
CC   infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and
CC   dense deposit disease. The method is useful in preventing or treating
CC   specific proliferative and oncogenic disease which includes sarcomas and
CC   carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,
CC   fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,
CC   hypoproliferative disorders, physical trauma, lesions and wounds. The
CC   method is also used in gene therapy. The present sequence is human Stat3
CC   protein
XX
SQ   Sequence 769 AA;

  Query Match             100.0%;  Score 494;  DB 5;  Length 769;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-54;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              |||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 9
AAE22056
ID   AAE22056 standard; protein; 769 AA.
XX
AC   AAE22056;
XX
DT   25-JUL-2002  (first entry)
XX
DE   Human protein related to angiogenesis regulation.
XX
KW   Human; signal transducer and activator of transcription 3; ischaemia;
KW   immune response; Stat3; coronary atherosclerosis; vascular occlusion;
KW   hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;
KW   inflammation; chronic obstructive pulmonary disease; cardiac arrest;
KW   insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;
KW   shock; chronic active hepatitis; adult respiratory distress syndrome;
KW   nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;
KW   Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;
KW   polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;
KW   proliferative disease; Grave's disease; ulcerative colitis; sarcoma;
KW   carcinoma; degenerative disorder; gene therapy; growth deficiency;
KW   cirrhosis; hypoproliferative disorder; lesion.
XX
OS   Homo sapiens.
XX
PN   WO200220032-A1.
XX
PD   14-MAR-2002.
XX
PF   10-SEP-2001; 2001WO-US028254.
XX
PR   08-SEP-2000; 2000US-0231212P.
XX
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (UYSF-) UNIV SOUTH FLORIDA.
XX
PI   Yu H,  Pardoll D,  Jove R,  Dalton W;
XX
DR   WPI; 2002-362218/39.
XX
PT   Modulating angiogenesis and an immune response in an individual, for
PT   treating a hypoxic or ischemic condition, comprises administering a
PT   compound that modulates the activity of a signal transducer and activator
```

PT of transcription 3.
XX
PS Disclosure; Page 83-85; 94pp; English.
XX
CC The invention relates to a method of modulating angiogenesis and immune
CC response. Method involves administering to an individual a compound that
CC modulate the activity of signal transducer and activator of transcription
CC 3 (Stat3). Modulating angiogenesis is useful for treating or preventing
CC hypoxic or ischaemic condition or disorder which is the result of stroke,
CC ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,
CC tissue ischaemia in the lower extremities, infarction, trauma, vascular
CC occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,
CC chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,
CC epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest,
CC nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy
CC with neovascularisation. Suppressing an immune response is useful for
CC ameliorating a symptom of an autoimmune disease such as systemic lupus
CC erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,
CC Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,
CC mixed connective tissue disease, primary biliary cirrhosis, pernicious
CC anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,
CC gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,
CC idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's
CC disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune
CC infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and
CC dense deposit disease. The method is useful in preventing or treating
CC specific proliferative and oncogenic disease which includes sarcomas and
CC carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,
CC fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,
CC hypoproliferative disorders, physical trauma, lesions and wounds. The
CC method is also used in gene therapy. The present sequence is human
CC protein related to angiogenesis regulation
XX
SQ Sequence 769 AA;

Query Match 100.0%; Score 494; DB 5; Length 769;
Best Local Similarity 100.0%; Pred. No. 5.3e-54;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
Db 282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy 61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
Db 342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377

RESULT 10
AAR72082
ID AAR72082 standard; protein; 770 AA.
XX
AC AAR72082;
XX
DT 25-MAR-2003 (revised)
DT 27-SEP-1995 (first entry)
XX
DE Mouse Stat3 (19sf6).
XX
KW Signal transducer and activator of transcription; STAT; 19sf6; Stat3;
KW receptor recognition factor; transcription factor; cellular debilitation;
KW derangement; dysfunction; interferon-gamma.
XX
OS Mus sp.
XX
PN WO9508629-A1.
XX
PD 30-MAR-1995.
XX
PF 26-SEP-1994; 94WO-US010849.
XX
PR 24-SEP-1993; 93US-00126588.
PR 24-SEP-1993; 93US-00126595.
PR 11-MAR-1994; 94US-00212184.
PR 11-MAR-1994; 94US-00212185.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI Darnell JE, Schindler CW, Shuai K, Wen Z, Zhong Z;
XX
DR WPI; 1995-139598/18.
DR N-PSDB; AAQ89340.
XX
PT Receptor recognition factor implicated in transcriptional stimulation of
PT genes - useful in drug screening assays and/or for treating cellular
PT debilitations, derangements and/or dysfunctions, etc.
XX
PS Claim 1; Page 107-110; 160pp; English.
XX
CC A fragment encoding the human Stat91 protein was used to screen a murine
CC thymus and spleen cDNA for homologous proteins. A highly homologous gene
CC (given in AAQ89338) was isolated that encoded a 91 kDa protein (AAR72080)
CC (Stat1) that was responsive to interferon- gamma. Using a fragment of the
CC mouse gene as probe, 2 additional members of the 113-91 family of
CC receptor recognition factor proteins were isolated. The 2 genes (AAQ89339
CC -40) were cloned in plasmids 13sf1 and 19sf6 and encoded proteins termed
CC Stat4 (AAR72081) and Stat3 (AAR72082), respectively. (Updated on 25-MAR-
CC 2003 to correct PN field.)
XX
SQ Sequence 770 AA;

Query Match 100.0%; Score 494; DB 2; Length 770;
Best Local Similarity 100.0%; Pred. No. 5.3e-54;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
Db 282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy 61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
Db 342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377

RESULT 11
AAR82995
ID AAR82995 standard; protein; 770 AA.
XX
AC AAR82995;
XX
DT 25-MAR-1996 (first entry)
XX
DE Mouse liver acute phase response factor.
XX
KW Mouse; acute phase response factor; transcription factor; interleukin-6;
KW signal transmission; liver; antibody; antisense; ribozyme;
KW antiinflammatory; antitumor; hypotensive; therapy.
XX
OS Mus musculus.
XX
PN EP676469-A2.
XX
PD 11-OCT-1995.
XX
PF 29-MAR-1995; 95EP-00104670.
XX
PR 04-APR-1994; 94JP-00065825.
XX
PA (KISH/) KISHIMOTO T.
XX
PI Akira S, Kishimoto T;
XX
DR WPI; 1995-346089/45.
DR N-PSDB; AAT05619.
XX
PT New acute phase response factor - for developing inhibitory agents for

```
PT   treating diseases induced by cytokine(s) such as IL-6, e.g. inflammatory
PT   diseases.
XX
PS   Claim 10; Page 20-22; 31pp; English.
XX
CC   The sequence represents a mouse acute phase response factor (APRF), a
CC   transcription factor related to signal transmission of interleukin-6 (IL-
CC   6). The protein is encoded by a cDNA, isolated from a mouse liver cDNA
CC   library using a polymerase chain reaction product (amplified using
CC   primers derived from an IL-6-treated mouse liver peptide) as a probe.
CC   APRF-inhibitors, e.g. antibodies, antisense oligonucleotides or
CC   ribozymes, may be used to treat diseases induced by IL-6, e.g.
CC   inflammatory disease, leukemia, cancer, osteoclasia, pulmonary
CC   hypertension, etc
XX
SQ   Sequence 770 AA;

  Query Match             100.0%;  Score 494;  DB 2;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-54;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 12
AAW03176
ID   AAW03176 standard; protein; 770 AA.
XX
AC   AAW03176;
XX
DT   24-OCT-1996  (first entry)
XX
DE   Mouse STAT4.
XX
KW   STAT; STAT4; signal transducer and activator of transcription;
KW   DNA binding protein; ligand; receptor; oncogenesis; inflammation;
KW   autoimmune disease; antagonist; therapy.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Domain          398. .508
FT                   /label= DNA_binding_domain
FT                   /note= "Claim 3, page 110"
XX
PN   WO9620954-A2.
XX
PD   11-JUL-1996.
XX
PF   28-DEC-1995;   95WO-US017025.
XX
PR   06-JAN-1995;   95US-00369796.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Darnell JE,  Wen Z,  Horvath CM,  Zhong Z;
XX
DR   WPI; 1996-333941/33.
DR   N-PSDB; AAT31280.
XX
PT   New STAT protein DNA-binding domain peptide(s) - useful for diagnosing,
PT   preventing or treating cellular dysfunction, e.g. oncogenesis,
PT   inflammation, parasitic disease or autoimmunity.
XX
PS   Disclosure; Page 87-90; 138pp; English.
XX
CC   Mouse signal transducer and activator of transcription (STAT) protein
CC   STAT4 (AAW03176) serves a dual purpose, i.e. signal transduction from
CC   ligand-activated receptor kinase complexes followed by nuclear
CC   translocation and DNA binding to activate transcription. Recombinant
CC   STAT4 can be obtd. using cDNA clone 19sf6 (AAT31278) obtd. from
CC   splenic/thymic cells. STAT4 includes a DNA-binding domain (see also
CC   AAW03167) capable of both receptor recognition and message delivery via
CC   DNA binding in a receptor-ligand specific manner. STAT proteins and their
CC   DNA binding domains (see also AAW03165-75) are useful for screening
CC   antagonists used to inhibit STAT-mediated signal transduction and
CC   activation of transcription
XX
SQ   Sequence 770 AA;

  Query Match             100.0%;  Score 494;  DB 2;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-54;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 13
AAY03768
ID   AAY03768 standard; protein; 770 AA.
XX
AC   AAY03768;
XX
DT   11-JUN-1999  (first entry)
XX
DE   Human STAT3 allelic variant.
XX
KW   Signal Transducer and Activator of Transcription 3; STAT3; allele; IL-6;
KW   intracellular transcription factor; interleukin-6; medicament; variant;
KW   pharmaceutical; autoimmune disease; inflammatory; human.
XX
OS   Homo sapiens.
XX
PN   EP905234-A2.
XX
PD   31-MAR-1999.
XX
PF   18-FEB-1998;   98EP-00102774.
XX
PR   16-SEP-1997;   97EP-00116061.
XX
PA   (ISTF ) ARS APPLIED RES SYSTEMS HOLDING NV.
XX
PI   Serlupi-Crescenzi O,  Della Pietra L;
XX
DR   WPI; 1999-192664/17.
DR   N-PSDB; AAX29281.
XX
PT   New human Signal Transducer and Activator of Transcription 3 (STAT3)
PT   allelic variant useful for treatment of autoimmune and inflammatory
PT   disease.
XX
PS   Claim 2; Page 9-13; 32pp; English.
XX
CC   The present sequence represents a predominant allelic variant of human
CC   Signal Transducer and Activator of Transcription 3 (STAT3) protein, an
CC   intracellular transcription factor which mediates IL-6 signals. The
CC   encoding sequence differs from the original published human STAT3 gene
CC   sequence (Akira et al., 1994). Host cells comprising a recombinant STAT3
CC   DNA molecule can be used for the recombinant expression of the variant.
CC   STAT3 protein is useful as a medicament or pharmaceutical composition for
CC   treatment of autoimmune or inflammatory diseases
XX
SQ   Sequence 770 AA;
```

```
  Query Match             100.0%;  Score 494;  DB 2;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-54;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 14
AAB12377
ID   AAB12377 standard; peptide; 770 AA.
XX
AC   AAB12377;
XX
DT   08-NOV-2000  (first entry)
XX
DE   N-terminal domain of murine STAT-3 protein.
XX
KW   STAT; signal transducer and activator of transcription; crystal;
KW   drug design; murine.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Region          4. .9
FT                   /label= Alpha helix 1
FT   Region          12. .21
FT                   /label= Alpha helix 2
FT   Region          19. .21
FT                   /label= 3(10) helix of alpha helix 2
FT   Region          28. .33
FT                   /label= Alpha helix 3
FT   Region          35. .40
FT                   /label= Alpha helix 4
FT   Region          43. .47
FT                   /label= Alpha helix 5
FT   Region          50. .73
FT                   /label= Alpha helix 6
FT   Region          77. .96
FT                   /label= Alpha helix 7
FT   Region          99. .119
FT                   /label= Alpha helix 8
XX
PN   US6087478-A.
XX
PD   11-JUL-2000.
XX
PF   23-JAN-1998;   98US-00012710.
XX
PR   23-JAN-1998;   98US-00012710.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Vinkemeier U,  Moarefi I,  Darnell JE,  Kuriyan J;
XX
DR   WPI; 2000-505108/45.
XX
PT   New crystals of an N-terminal fragment of a signal transducer and
PT   activator of transcription that effectively diffracts x-rays, useful for
PT   drug screening and development.
XX
PS   Disclosure; Fig 1; 42pp; English.
XX
CC   The present invention relates to a crystal of an N-terminal fragment of a
CC   signal transducer and activator of transcription (STAT) protein. The
CC   crystal effectively diffracts X-rays, allowing the determination of the
CC   atomic coordinates of the N-terminal domain to a resolution of greater
CC   than 5.0 Angstroms. The present sequence is the N-terminal domain of the
CC   murine STAT 3 protein. The N-terminal domain enables STAT dimers to
CC   interact and bind DNA cooperatively, a mechanism important for gene
CC   activation. The crystals are useful in drug screening and development by
CC   selecting a potential drug by performing rational drug design with the 3-
CC   dimensional structure determined for the crystal
XX
SQ   Sequence 770 AA;

  Query Match             100.0%;  Score 494;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 5.3e-54;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 15
AAE14652
ID   AAE14652 standard; protein; 770 AA.
XX
AC   AAE14652;
XX
DT   16-JUL-2002  (first entry)
XX
DE   Murine STAT3 protein.
XX
KW   Signal transducer and activator of transcription; STAT3;
KW   drug development; drug discovery; crystal; inflammation; allergy; asthma;
KW   leukaemia; anaemia; neutropaenia; thrombocytopaenia; cancer; obesity;
KW   viral disease; growth retardation; murine.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Domain          1. .130
FT                   /note= "Conserved N-terminal domain of the STAT family"
XX
PN   US6312887-B1.
XX
PD   06-NOV-2001.
XX
PF   24-APR-2000; 2000US-00556273.
XX
PR   23-JAN-1998;   98US-00012710.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Vinkemeier U,  Moarefi I,  Darnell JE,  Kuriyan J;
XX
DR   WPI; 2002-033337/04.
XX
PT   Identifying compounds that bind to signal transducer and activator of
PT   transcription proteins, useful for the production of new drugs.
XX
PS   Example; Col 47-50; 44pp; English.
XX
CC   The invention relates to methods for detecting compounds that bind to
CC   signal transducer and activator of transcription (STAT) proteins for the
CC   discovery and development of new drug compounds based on the structural
CC   properties of the protein crystal. The methods include: identifying a
CC   compound that binds to the N-terminal domain of a STAT protein,
CC   identifying a compound that enhances or diminishes the binding of the
CC   dimeric STAT proteins to each other and/or their nucleic acid binding
CC   site; or identifying a compound that enhances or diminishes the ability
CC   of STAT protein dimers to induce the expression of a gene operably under
CC   the control of a promoter containing at least two adjacent weak binding
CC   sites for STAT protein dimers. The methods are used for identifying new
```

```
CC   drugs. An antagonist of STAT N-terminal dimeric interactions that
CC   inhibits the binding of the STAT dimers to adjacent weak binding sites on
CC   a promoter of a gene, could be useful as drugs in the treatment of
CC   diseases, e.g. inflammation, allergy, asthma and leukaemias. On the other
CC   hand, an agonist of N-terminal dimeric interactions between STAT dimers,
CC   can be used as drugs in the treatment of diseases e.g. anaemia,
CC   neutropaenia, thrombocytopaenia, cancer, obesity, viral diseases and
CC   growth retardation. The present sequence is murine STAT3 protein
XX
SQ   Sequence 770 AA;

  Query Match            100.0%;  Score 494;  DB 5;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 5.3e-54;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


Search completed: May 25, 2005, 17:36:36
Job time : 51.265 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:25:01 ; Search time 12.0119 Seconds
(without alignments)
596.600 Million cell updates/sec

Title: US-10-090-185-25
Perfect score: 494
Sequence: 1 KKLEELQQKVSYKGDPIVQH.....LNYQLKIKVCIDKDSGDVAA 96

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	494	100.0	96	3	US-09-387-418A-25	Sequence 25, Appl
2	494	100.0	129	3	US-09-387-418A-24	Sequence 24, Appl
3	494	100.0	185	3	US-09-387-418A-23	Sequence 23, Appl
4	494	100.0	223	3	US-09-387-418A-22	Sequence 22, Appl
5	494	100.0	271	3	US-09-387-418A-9	Sequence 9, Appli
6	494	100.0	770	1	US-08-369-796-12	Sequence 12, Appl
7	494	100.0	770	1	US-08-416-581B-9	Sequence 9, Appli
8	494	100.0	770	2	US-08-852-091-12	Sequence 12, Appl
9	494	100.0	770	2	US-08-820-754-12	Sequence 12, Appl
10	494	100.0	770	3	US-08-956-652-12	Sequence 12, Appl
11	494	100.0	770	3	US-08-956-869-12	Sequence 12, Appl
12	494	100.0	770	3	US-09-012-710-8	Sequence 8, Appli
13	494	100.0	770	3	US-08-948-547-12	Sequence 12, Appl
14	494	100.0	770	3	US-09-364-970-3	Sequence 3, Appli
15	494	100.0	770	3	US-09-364-970-5	Sequence 5, Appli
16	494	100.0	770	3	US-09-556-273-8	Sequence 8, Appli
17	494	100.0	770	3	US-08-956-653A-12	Sequence 12, Appl
18	494	100.0	770	3	US-09-526-542-2	Sequence 2, Appli
19	494	100.0	770	4	US-08-212-185-12	Sequence 12, Appl
20	494	100.0	770	4	US-10-117-087-2	Sequence 2, Appli
21	494	100.0	770	5	PCT-US95-17025-12	Sequence 12, Appl
22	489	99.0	770	1	US-08-416-581B-1	Sequence 1, Appli
23	489	99.0	770	1	US-08-416-581B-5	Sequence 5, Appli
24	489	99.0	770	3	US-09-087-465-6	Sequence 6, Appli
25	489	99.0	770	4	US-09-972-800A-6	Sequence 6, Appli
26	489	99.0	771	1	US-08-276-099A-14	Sequence 14, Appl
27	489	99.0	771	1	US-08-781-890-14	Sequence 14, Appl
28	396	80.2	229	3	US-09-387-418A-18	Sequence 18, Appl
29	396	80.2	229	3	US-09-387-418A-28	Sequence 28, Appl
30	396	80.2	229	3	US-09-387-418A-30	Sequence 30, Appl
31	396	80.2	229	3	US-09-387-418A-31	Sequence 31, Appl
32	396	80.2	252	3	US-09-387-418A-14	Sequence 14, Appl
33	379	76.7	229	3	US-09-387-418A-29	Sequence 29, Appl
34	318	64.4	213	3	US-09-387-418A-19	Sequence 19, Appl
35	318	64.4	236	3	US-09-387-418A-15	Sequence 15, Appl
36	310	62.8	268	3	US-09-387-418A-12	Sequence 12, Appl
37	310	62.8	582	4	US-09-430-806A-3	Sequence 3, Appli
38	310	62.8	712	1	US-08-369-796-6	Sequence 6, Appli
39	310	62.8	712	2	US-08-852-091-6	Sequence 6, Appli
40	310	62.8	712	2	US-08-820-754-6	Sequence 6, Appli
41	310	62.8	712	3	US-08-956-652-6	Sequence 6, Appli
42	310	62.8	712	3	US-08-956-869-6	Sequence 6, Appli
43	310	62.8	712	3	US-08-948-547-6	Sequence 6, Appli
44	310	62.8	712	3	US-08-956-653A-6	Sequence 6, Appli
45	310	62.8	712	4	US-08-212-185-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-09-387-418A-25
; Sequence 25, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/09/387,418A
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-387-418A-25

Query Match 100.0%; Score 494; DB 3; Length 96;
Best Local Similarity 100.0%; Pred. No. 4.3e-54;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
```

RESULT 2
US-09-387-418A-24
; Sequence 24, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/09/387,418A
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43

```
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 24
;   LENGTH: 129
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-24

  Query Match            100.0%;  Score 494;  DB 3;  Length 129;
  Best Local Similarity   100.0%;  Pred. No. 6.3e-54;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         34 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 93

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db         94 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 129


RESULT 3
US-09-387-418A-23
; Sequence 23, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
;  APPLICANT:  Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 23
;   LENGTH: 185
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-23

  Query Match            100.0%;  Score 494;  DB 3;  Length 185;
  Best Local Similarity   100.0%;  Pred. No. 1e-53;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 149

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        150 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 185


RESULT 4
US-09-387-418A-22
; Sequence 22, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
;  APPLICANT:  Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 22
;   LENGTH: 223
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-22

  Query Match            100.0%;  Score 494;  DB 3;  Length 223;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-53;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 187

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        188 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 223


RESULT 5
US-09-387-418A-9
; Sequence 9, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
;  APPLICANT:  Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 271
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-9

  Query Match            100.0%;  Score 494;  DB 3;  Length 271;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-53;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 235

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        236 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271


RESULT 6
US-08-369-796-12
; Sequence 12, Application US/08369796
; Patent No. 5716622
;  GENERAL INFORMATION:
;    APPLICANT:  James E. Darnell, Jr.
;    APPLICANT:  Zilong Wen
;    APPLICANT:  Curt M. Horvath
;    APPLICANT:  Zhong Zhong
;    TITLE OF INVENTION:  FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
;    TITLE OF INVENTION:  TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
```

```
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/369,796
;       FILING DATE:  06-JAN-1995
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Jackson Esq., David A.
;       REGISTRATION NUMBER:  26,742
;       REFERENCE/DOCKET NUMBER:  600-1-116
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  201 487-5800
;       TELEFAX:  201 343-1684
;       TELEX:  133521
;   INFORMATION FOR SEQ ID NO:  12:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  770 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-369-796-12

  Query Match             100.0%;  Score 494;  DB 1;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-53;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 7
US-08-416-581B-9
; Sequence 9, Application US/08416581B
; Patent No. 5719042
;  GENERAL INFORMATION:
;    APPLICANT:  Kishimoto, Tadamitsu
;    APPLICANT:  Akira, Shizuo
;    TITLE OF INVENTION:  TRANSCRIPTION FACTOR APRF
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SUGHRUE, MION, ZINN, MACPEAK & SEAS
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/416,581B
;      FILING DATE:  04-APR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-65825/1994
;      FILING DATE:  04-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Nakamura, Dean H.
;      REGISTRATION NUMBER:  33,981
;      REFERENCE/DOCKET NUMBER:  Q-37891
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)293-7060
;      TELEFAX:  (202)293-7860
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-416-581B-9

  Query Match             100.0%;  Score 494;  DB 1;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-53;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 8
US-08-852-091-12
; Sequence 12, Application US/08852091
; Patent No. 5883228
;  GENERAL INFORMATION:
;    APPLICANT:  James E. Darnell, Jr.
;    APPLICANT:  Zilong Wen
;    APPLICANT:  Curt M. Horvath
;    APPLICANT:  Zhong Zhong
;    TITLE OF INVENTION:  FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
;    TITLE OF INVENTION:  TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/852,091
;      FILING DATE:  06-MAY-1997
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/369,796
;      FILING DATE:  06-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-116
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-852-091-12

  Query Match             100.0%;  Score 494;  DB 2;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-53;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341
```

```
Qy          61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
               |||||||||||||||||||||||||||||||||||||
Db         342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377
```

```
RESULT 9
US-08-820-754-12
; Sequence 12, Application US/08820754
; Patent No. 5976835
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/820,754
;      FILING DATE:  19-MAR-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,185
;      FILING DATE:  11-MAR-1994
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-820-754-12

  Query Match             100.0%;  Score 494;  DB 2;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-53;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341
```

```
Qy          61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
               |||||||||||||||||||||||||||||||||||||
Db         342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377
```

```
RESULT 10
US-08-956-652-12
; Sequence 12, Application US/08956652
; Patent No. 6013475
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/956,652
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,185
;      FILING DATE:  11-MAR-1994
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-956-652-12

  Query Match             100.0%;  Score 494;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-53;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy           1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy          61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
```

```
                                     ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 11
US-08-956-869-12
; Sequence 12, Application US/08956869
; Patent No. 6030808
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/956,869
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/212,185
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-956-869-12

  Query Match             100.0%;  Score 494;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-53;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 12
US-09-012-710-8
; Sequence 8, Application US/09012710
; Patent No. 6087478
;  GENERAL INFORMATION:
;    APPLICANT:  Vinkemeier, Uwe
;    APPLICANT:  Moarefi, Ismail
;    APPLICANT:  Darnell, Jr., James E.
;    APPLICANT:  Kuriyan, John
;    TITLE OF INVENTION:  A CRYSTAL OF THE N-TERMINAL DOMAIN OF A
;    TITLE OF INVENTION:  STAT PROTEIN AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/012,710
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-194
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
;      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-09-012-710-8

  Query Match             100.0%;  Score 494;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-53;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 13
US-08-948-547-12
; Sequence 12, Application US/08948547
; Patent No. 6124118
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
```

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/948,547
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/212,185
;      FILING DATE:  11-MAR-1994
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-948-547-12

  Query Match             100.0%;  Score 494;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-53;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 14
US-09-364-970-3
; Sequence 3, Application US/09364970
; Patent No. 6235873
; GENERAL INFORMATION:
;  APPLICANT: Bromberg, Jacqueline
;  TITLE OF INVENTION: CONSTITUTIVELY ACTIVE TRANSCRIPTION FACTORS AND THEIR
;  TITLE OF INVENTION: USES FOR IDENTIFYING MODULATORS OF ACTIVITY INCLUDING
;  TITLE OF INVENTION: DYSPROLIFERATIVE CELLULAR CHANGES
;  FILE REFERENCE: 600-1-252
;  CURRENT APPLICATION NUMBER: US/09/364,970
;  CURRENT FILING DATE: 1999-07-31
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-364-970-3

  Query Match             100.0%;  Score 494;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-53;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 15
US-09-364-970-5
; Sequence 5, Application US/09364970
; Patent No. 6235873
; GENERAL INFORMATION:
;  APPLICANT: Bromberg, Jacqueline
;  TITLE OF INVENTION: CONSTITUTIVELY ACTIVE TRANSCRIPTION FACTORS AND THEIR
;  TITLE OF INVENTION: USES FOR IDENTIFYING MODULATORS OF ACTIVITY INCLUDING
;  TITLE OF INVENTION: DYSPROLIFERATIVE CELLULAR CHANGES
;  FILE REFERENCE: 600-1-252
;  CURRENT APPLICATION NUMBER: US/09/364,970
;  CURRENT FILING DATE: 1999-07-31
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-364-970-5

  Query Match             100.0%;  Score 494;  DB 3;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-53;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


Search completed: May 25, 2005, 17:47:39
Job time : 13.0119 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:36:58 ; Search time 42.0417 Seconds
(without alignments)
763.830 Million cell updates/sec

Title: US-10-090-185-25
Perfect score: 494
Sequence: 1 KKLEELQQKVSYKGDPIVQH.....LNYQLKIKVCIDKDSGDVAA 96

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1434725 seqs, 334507595 residues

Total number of hits satisfying chosen parameters: 1434725

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	494	100.0	96	13	US-10-090-185-25	Sequence 25, Appl
2	494	100.0	129	13	US-10-090-185-24	Sequence 24, Appl
3	494	100.0	185	13	US-10-090-185-23	Sequence 23, Appl
4	494	100.0	223	13	US-10-090-185-22	Sequence 22, Appl
5	494	100.0	271	13	US-10-090-185-9	Sequence 9, Appli
6	494	100.0	720	15	US-10-380-020-4	Sequence 4, Appli
7	494	100.0	769	15	US-10-380-020-2	Sequence 2, Appli
8	494	100.0	769	15	US-10-380-020-5	Sequence 5, Appli
9	494	100.0	770	11	US-09-876-773-12	Sequence 12, Appl
10	494	100.0	770	14	US-10-045-792-8	Sequence 8, Appli
11	494	100.0	770	14	US-10-038-010-56	Sequence 56, Appl
12	494	100.0	770	14	US-10-117-087-2	Sequence 2, Appli
13	494	100.0	770	17	US-10-639-617-12	Sequence 12, Appl
14	494	100.0	793	9	US-09-925-302-780	Sequence 780, App
15	494	100.0	793	10	US-09-925-302-780	Sequence 780, App
16	491	99.4	770	15	US-10-116-275-329	Sequence 329, App
17	489	99.0	770	15	US-10-116-275-349	Sequence 349, App
18	396	80.2	229	13	US-10-090-185-18	Sequence 18, Appl
19	396	80.2	229	13	US-10-090-185-28	Sequence 28, Appl
20	396	80.2	229	13	US-10-090-185-30	Sequence 30, Appl
21	396	80.2	229	13	US-10-090-185-31	Sequence 31, Appl
22	396	80.2	252	13	US-10-090-185-14	Sequence 14, Appl
23	379	76.7	229	13	US-10-090-185-29	Sequence 29, Appl
24	318	64.4	213	13	US-10-090-185-19	Sequence 19, Appl
25	318	64.4	236	13	US-10-090-185-15	Sequence 15, Appl
26	311	63.0	749	9	US-09-833-205-4	Sequence 4, Appli
27	310	62.8	268	13	US-10-090-185-12	Sequence 12, Appl
28	310	62.8	582	14	US-10-245-120-3	Sequence 3, Appli
29	310	62.8	712	11	US-09-876-773-6	Sequence 6, Appli
30	310	62.8	712	14	US-10-245-120-2	Sequence 2, Appli
31	310	62.8	712	17	US-10-639-617-6	Sequence 6, Appli
32	310	62.8	712	17	US-10-936-390-5	Sequence 5, Appli
33	310	62.8	750	9	US-09-833-205-2	Sequence 2, Appli
34	310	62.8	750	11	US-09-876-773-4	Sequence 4, Appli
35	310	62.8	750	14	US-10-245-120-1	Sequence 1, Appli
36	310	62.8	750	14	US-10-308-279-44	Sequence 44, Appl
37	310	62.8	750	16	US-10-755-889-352	Sequence 352, App
38	310	62.8	750	16	US-10-755-889-823	Sequence 823, App
39	310	62.8	750	17	US-10-492-043-19	Sequence 19, Appl
40	310	62.8	750	17	US-10-639-617-4	Sequence 4, Appli
41	310	62.8	786	9	US-09-925-297-550	Sequence 550, App
42	291	58.9	423	14	US-10-177-293-437	Sequence 437, App
43	291	58.9	748	14	US-10-177-293-439	Sequence 439, App
44	291	58.9	748	16	US-10-755-889-70	Sequence 70, Appl
45	291	58.9	748	16	US-10-755-889-362	Sequence 362, App

ALIGNMENTS

RESULT 1
US-10-090-185-25
; Sequence 25, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/10/090,185
; CURRENT FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: 09/387,418
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-090-185-25

Query Match 100.0%; Score 494; DB 13; Length 96;
Best Local Similarity 100.0%; Pred. No. 2.9e-53;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
```

```
RESULT 2
US-10-090-185-24
; Sequence 24, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR F
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
;   LENGTH: 129
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-24

  Query Match            100.0%;  Score 494;  DB 13;  Length 129;
  Best Local Similarity  100.0%;  Pred. No. 4.2e-53;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         34 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 93

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db         94 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 129


RESULT 3
US-10-090-185-23
; Sequence 23, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR F
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
;   LENGTH: 185
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-23

  Query Match            100.0%;  Score 494;  DB 13;  Length 185;
  Best Local Similarity  100.0%;  Pred. No. 6.7e-53;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 149

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        150 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 185


RESULT 4
US-10-090-185-22
; Sequence 22, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
;   LENGTH: 223
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-22

  Query Match            100.0%;  Score 494;  DB 13;  Length 223;
  Best Local Similarity  100.0%;  Pred. No. 8.5e-53;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 187

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        188 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 223


RESULT 5
US-10-090-185-9
; Sequence 9, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 271
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-9

  Query Match            100.0%;  Score 494;  DB 13;  Length 271;
  Best Local Similarity  100.0%;  Pred. No. 1.1e-52;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 235

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
```

```
Db        236 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 271


RESULT 6
US-10-380-020-4
; Sequence 4, Application US/10380020
; Publication No. US20040052762A1
; GENERAL INFORMATION:
;  APPLICANT: Yu, Hua
;  APPLICANT:  Pardoll, Drew
;  APPLICANT:  Jove, Richard
;  APPLICANT:  Dalton, Willian
;  TITLE OF INVENTION: Stat3 Agonists and Antagonists and Therapeutic Uses Thereof
;  FILE REFERENCE: 10873-009-999
;  CURRENT APPLICATION NUMBER: US/10/380,020
;  CURRENT FILING DATE:  2003-03-07
;  PRIOR APPLICATION NUMBER: 60/231,212
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
;   LENGTH: 720
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-380-020-4

  Query Match            100.0%;  Score 494;  DB 15;  Length 720;
  Best Local Similarity  100.0%;  Pred. No. 3.8e-52;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 7
US-10-380-020-2
; Sequence 2, Application US/10380020
; Publication No. US20040052762A1
; GENERAL INFORMATION:
;  APPLICANT: Yu, Hua
;  APPLICANT:  Pardoll, Drew
;  APPLICANT:  Jove, Richard
;  APPLICANT:  Dalton, Willian
;  TITLE OF INVENTION: Stat3 Agonists and Antagonists and Therapeutic Uses Thereof
;  FILE REFERENCE: 10873-009-999
;  CURRENT APPLICATION NUMBER: US/10/380,020
;  CURRENT FILING DATE:  2003-03-07
;  PRIOR APPLICATION NUMBER: 60/231,212
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 769
;   TYPE: PRT<
;   ORGANISM: Homo Sapiens
US-10-380-020-2

  Query Match            100.0%;  Score 494;  DB 15;  Length 769;
  Best Local Similarity  100.0%;  Pred. No. 4.2e-52;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 8
US-10-380-020-5
; Sequence 5, Application US/10380020
; Publication No. US20040052762A1
; GENERAL INFORMATION:
;  APPLICANT: Yu, Hua
;  APPLICANT:  Pardoll, Drew
;  APPLICANT:  Jove, Richard
;  APPLICANT:  Dalton, Willian
;  TITLE OF INVENTION: Stat3 Agonists and Antagonists and Therapeutic Uses Thereof
;  FILE REFERENCE: 10873-009-999
;  CURRENT APPLICATION NUMBER: US/10/380,020
;  CURRENT FILING DATE:  2003-03-07
;  PRIOR APPLICATION NUMBER: 60/231,212
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 769
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-380-020-5

  Query Match            100.0%;  Score 494;  DB 15;  Length 769;
  Best Local Similarity  100.0%;  Pred. No. 4.2e-52;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 9
US-09-876-773-12
; Sequence 12, Application US/09876773
; Publication No. US20040058318A1
;   GENERAL INFORMATION:
;        APPLICANT: Darnell Jr., James E.
;                   Schindler, Christian W.
;                   Fu, Xian-Yuan
;                   Wen, Zilong
;                   Zhong, Zhong
;        TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
;                            SEQUENCES AND METHODS OF USE THEREOF
;        NUMBER OF SEQUENCES: 25
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Klauber & Jackson
;             STREET: 411 Hackensack Avenue
;             CITY: Hackensack
;             STATE: New Jersey
;             COUNTRY: USA
;             ZIP: 07601
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/876,773
;             FILING DATE: 07-Jun-2001
;             CLASSIFICATION: 435
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/212,185
;             FILING DATE: 11-MAR-1994
;             APPLICATION NUMBER: US 07/980,498
;             FILING DATE: 23-NOV-1992
;             APPLICATION NUMBER: US 07/854,296
```

```
;              FILING DATE: 19-MAR-1992
;              APPLICATION NUMBER: WO US93/02569
;              FILING DATE: 19-MAR-1993
;              APPLICATION NUMBER: US 08/126,588
;              FILING DATE: 24-SEP-1993
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Jackson Esq., David A.
;              REGISTRATION NUMBER: 26,742
;              REFERENCE/DOCKET NUMBER: 600-1-073 CIP
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 201 487-5800
;              TELEFAX: 201 343-1684
;              TELEX: 133521
;    INFORMATION FOR SEQ ID NO: 12:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 770 amino acids
;              TYPE: amino acid
;              TOPOLOGY: linear
;         MOLECULE TYPE: protein
;         SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-876-773-12

  Query Match             100.0%;  Score 494;  DB 11;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-52;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 10
US-10-045-792-8
; Sequence 8, Application US/10045792
; Publication No. US20030003563A1
;   GENERAL INFORMATION:
;         APPLICANT: Vinkemeier, Uwe
;                    Moarefi, Ismail
;                    Darnell, Jr., James E.
;                    Kuriyan, John
;         TITLE OF INVENTION: A CRYSTAL OF THE N-TERMINAL DOMAIN OF A
;                             STAT PROTEIN AND METHODS OF USE THEREOF
;         NUMBER OF SEQUENCES: 13
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: Klauber & Jackson
;              STREET: 411 Hackensack Avenue, 4th Floor
;              CITY: Hackensack
;              STATE: New Jersey
;              COUNTRY: USA
;              ZIP: 07601
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: IBM PC compatible
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: PatentIn Release #1.0, Version #1.30
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/10/045,792
;              FILING DATE: 19-Oct-2001
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: US/09/012,710
;              FILING DATE: <Unknown>
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Jackson Esq., David A.
;              REGISTRATION NUMBER: 26,742
;              REFERENCE/DOCKET NUMBER: 600-1-194
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: 201-487-5800
;              TELEFAX: 201-343-1684
;              TELEX: 133521
;    INFORMATION FOR SEQ ID NO: 8:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 770 amino acids
;              TYPE: amino acid
;              STRANDEDNESS: single
;              TOPOLOGY: linear
;         MOLECULE TYPE: protein
;         HYPOTHETICAL: NO
;         SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-045-792-8

  Query Match             100.0%;  Score 494;  DB 14;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-52;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 11
US-10-038-010-56
; Sequence 56, Application US/10038010
; Publication No. US20030040089A1
; GENERAL INFORMATION:
;  APPLICANT: HYBRIGENICS
;  APPLICANT:  Pierre, Legrain
;  TITLE OF INVENTION: Protein-protein interactions in adipocyte cells
;  FILE REFERENCE: B4767A
;  CURRENT APPLICATION NUMBER: US/10/038,010
;  CURRENT FILING DATE:  2002-07-23
;  PRIOR APPLICATION NUMBER: US 60/259,377
;  PRIOR FILING DATE: 2001-01-02
;  NUMBER OF SEQ ID NOS: 67
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: STAT3 : Transcription factor
;   LOCATION: (1)..(770)
;   OTHER INFORMATION:
US-10-038-010-56

  Query Match             100.0%;  Score 494;  DB 14;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-52;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 12
US-10-117-087-2
; Sequence 2, Application US/10117087
; Publication No. US20030166854A1
; GENERAL INFORMATION:
;  APPLICANT: SERLUPI-CRESCENZI, Ottaviano
;  APPLICANT:  DELLA PIETRA, Linda
;  TITLE OF INVENTION: ALLELIC VARIANTS OF HUMAN STAT3
;  FILE REFERENCE: SERLUPI=2
;  CURRENT APPLICATION NUMBER: US/10/117,087
```

```
;  CURRENT FILING DATE:  2002-04-08
;  PRIOR APPLICATION NUMBER: US/09/526,542
;  PRIOR FILING DATE: 2000-03-19
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Human
US-10-117-087-2

  Query Match           100.0%;  Score 494;  DB 14;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 4.2e-52;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 13
US-10-639-617-12
; Sequence 12, Application US/10639617
; Publication No. US20050079543A1
;   GENERAL INFORMATION:
;        APPLICANT: Darnell Jr., James E.
;                   Schindler, Christian W.
;                   Fu, Xian-Yuan
;                   Wen, Zilong
;                   Zhong, Zhong
;        TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
;                            SEQUENCES AND METHODS OF USE THEREOF
;        NUMBER OF SEQUENCES: 25
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Klauber & Jackson
;             STREET: 411 Hackensack Avenue
;             CITY: Hackensack
;             STATE: New Jersey
;             COUNTRY: USA
;             ZIP: 07601
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/639,617
;             FILING DATE: 12-Aug-2003
;             CLASSIFICATION: 435
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/212,185
;             FILING DATE: 11-MAR-1994
;             APPLICATION NUMBER: US 07/980,498
;             FILING DATE: 23-NOV-1992
;             APPLICATION NUMBER: US 07/854,296
;             FILING DATE: 19-MAR-1992
;             APPLICATION NUMBER: WO US93/02569
;             FILING DATE: 19-MAR-1993
;             APPLICATION NUMBER: US 08/126,588
;             FILING DATE: 24-SEP-1993
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Jackson Esq., David A.
;             REGISTRATION NUMBER: 26,742
;             REFERENCE/DOCKET NUMBER: 600-1-073 CIP
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 201 487-5800
;             TELEFAX: 201 343-1684
;             TELEX: 133521
;   INFORMATION FOR SEQ ID NO: 12:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 770 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-639-617-12

  Query Match           100.0%;  Score 494;  DB 17;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 4.2e-52;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 14
US-09-925-302-780
; Sequence 780, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA104
;  CURRENT APPLICATION NUMBER: US/09/925,302
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05918
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 896
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 780
;   LENGTH: 793
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-925-302-780

  Query Match           100.0%;  Score 494;  DB 9;  Length 793;
  Best Local Similarity  100.0%;  Pred. No. 4.3e-52;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 364

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        365 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 400


RESULT 15
US-09-925-302-780
; Sequence 780, Application US/09925302
; Publication No. US20030064072A9
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA104
;  CURRENT APPLICATION NUMBER: US/09/925,302
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05918
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 896
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 780
```

; LENGTH: 793
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-302-780

  Query Match 100.0%; Score 494; DB 10; Length 793;
  Best Local Similarity 100.0%; Pred. No. 4.3e-52;
  Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 364

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        365 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 400
```

Search completed: May 25, 2005, 18:21:52
Job time : 43.375 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:24:07 ; Search time 9.0566 Seconds
(without alignments)
1019.898 Million cell updates/sec

Title: US-10-090-185-25
Perfect score: 494
Sequence: 1 KKLEELQQKVSYKGDPIVQH.....LNYQLKIKVCIDKDSGDVAA 96

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	494	100.0	770	2	I49508	ISGF3 p91-related
2	489	99.0	770	2	A54444	DNA-binding protei
3	310	62.8	739	2	A46159	interferon-depende
4	283	57.3	748	2	A56047	gamma-interferon a
5	204	41.3	851	2	A46160	interferon alpha-i
6	90.5	18.3	848	2	A54740	interleukin-4-indu
7	86.5	17.5	786	2	I49274	mammary gland fact
8	86.5	17.5	793	2	S54772	mammary gland fact
9	86.5	17.5	794	2	G02317	transcription acti
10	86.5	17.5	794	2	S55527	mammary gland fact
11	83.5	16.9	837	2	I57557	DNA-Binding Protei
12	72	14.6	393	2	S32458	4-hydroxyphenylpyr
13	72	14.6	393	2	S32821	4-hydroxyphenylpyr
14	71	14.4	498	2	T06667	argininosuccinate
15	68	13.8	363	2	S48483	MRS1 protein - yea
16	68	13.8	590	2	A48461	ovarian abundant m
17	68	13.8	1152	2	E71667	2-acylglycerophosp
18	67	13.6	376	2	S32820	alloantigen F - ra
19	67	13.6	402	2	T29703	hypothetical prote
20	67	13.6	415	2	JC7167	C kinase 1 interac
21	67	13.6	508	2	AF1479	fatty-acid-CoA lig
22	66	13.4	376	2	A60236	F protein - mouse
23	66	13.4	393	2	T25063	hypothetical prote
24	66	13.4	508	2	AC1119	fatty-acid-CoA lig
25	65.5	13.3	748	2	T47250	complex I intermed
26	65.5	13.3	926	2	T39082	hypothetical prote
27	64	13.0	334	1	DEBYMM	malate dehydrogena
28	64	13.0	364	2	T30084	hypothetical prote
29	64	13.0	460	2	H89907	hypothetical prote
30	63.5	12.9	395	2	T43892	translation elonga
31	63	12.8	416	2	A56486	perinuclear bindin
32	63	12.8	695	2	T39954	probable gtpase ac
33	62.5	12.7	339	2	T41126	meiotic recombinat
34	62.5	12.7	378	1	S42088	mevalonate kinase
35	62.5	12.7	781	2	T39373	dynamin-related pr
36	62.5	12.7	804	2	B84389	ribonucleoside red
37	62	12.6	307	1	PNBSL	beta-lactamase (EC
38	62	12.6	328	2	A99345	conserved hypothet
39	62	12.6	369	2	T21556	hypothetical prote
40	62	12.6	580	2	C71182	probable ABC trans
41	62	12.6	1493	2	A38218	GAP-associated pro
42	61.5	12.4	191	2	B70042	hypothetical prote
43	61.5	12.4	256	2	T48076	60S RIBOSOMAL PROT
44	61.5	12.4	451	2	S73433	pre-B cell enhanci
45	61.5	12.4	451	2	B81056	proteinase, probab

ALIGNMENTS

RESULT 1
I49508
ISGF3 p91-related transcription factor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I49508; I49009
R;Akira, S.; Nishio, Y.; Inoue, M.; Wang, X.J.; Wei, S.; Matsusaka, T.; Yoshida, K.; Su
Cell 77, 63-71, 1994
A;Title: Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related tr
A;Reference number: A54444; MUID:94208062; PMID:7512451
A;Accession: I49508
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-770 <RES>
A;Cross-references: UNIPROT:P42227; GB:L29278; NID:g476715; PIDN:AAA37254.1; PID:g47671
R;Raz, R.; Durbin, J.E.; Levy, D.E.
J. Biol. Chem. 269, 24391-24395, 1994
A;Title: Acute phase response factor and additional members of the interferon-stimulate
A;Reference number: I49009; MUID:95014185; PMID:7523373
A;Accession: I49009
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-393,'M',395-700,702-770 <RE2>
A;Cross-references: EMBL:U08378; NID:g473889; PIDN:AAA56668.1; PID:g473890
C;Genetics:
A;Gene: APRF
C;Superfamily: human signal transducer and transcription activator STAT5A

Query Match 100.0%; Score 494; DB 2; Length 770;
Best Local Similarity 100.0%; Pred. No. 4.6e-46;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377
```

RESULT 2
A54444
DNA-binding protein APRF - human
C;Species: Homo sapiens (man)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 09-Jul-2004
C;Accession: A54444
R;Akira, S.; Nishio, Y.; Inoue, M.; Wang, X.J.; Wei, S.; Matsusaka, T.; Yoshida, K.; Su
Cell 77, 63-71, 1994
A;Title: Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related tr
A;Reference number: A54444; MUID:94208062; PMID:7512451
A;Accession: A54444

```
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-770 <RES>
A;Cross-references: UNIPROT:P40763; GB:L29277; NID:g475788; PID:g475789
C;Genetics:
A;Gene: GDB:STAT3; APRF
A;Cross-references: GDB:358950
A;Map position: 17q21-17q21
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: DNA binding; transcription factor

  Query Match             99.0%;  Score 489;  DB 2;  Length 770;
  Best Local Similarity   99.0%;  Pred. No. 1.6e-45;
  Matches   95;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELHQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 3
A46159
interferon-dependent positive-acting transcription factor ISGF-3 91K chain - human
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A46159
R;Schindler, C.; Fu, X.Y.; Improta, T.; Aebersold, R.; Darnell Jr., J.E.
Proc. Natl. Acad. Sci. U.S.A. 89, 7836-7839, 1992
A;Title: Proteins of transcription factor ISGF-3: one gene encodes the 91-and 84-kDa ISG
A;Reference number: A46159; MUID:92366557; PMID:1502203
A;Accession: A46159
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid; protein
A;Residues: 1-739 <SCH>
A;Cross-references: UNIPROT:P42224
A;Experimental source: HeLa cells
A;Note: sequence extracted from NCBI backbone (NCBIP:110818)
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             62.8%;  Score 310;  DB 2;  Length 739;
  Best Local Similarity   64.4%;  Pred. No. 6.6e-26;
  Matches   58;  Conservative   15;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||:|| :|: ||| ::: :| :|   ||: |::|:||||||||||| || ||||:||
Db        267 KKLEELEQKYTYEHDPITKNKQVLWDRTFSLFQQLIQSSFVVERQPCMPTHPQRPLVLKT 326

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKD 90
              ||||| |:||||  |||| ||:||  |||
Db        327 GVQFTVKLRLLVKLQELNYNLKVKVLFDKD 356


RESULT 4
A56047
gamma-interferon activation site-binding protein Stat4 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004
C;Accession: A56047
R;Yamamoto, K.; Quelle, F.W.; Thierfelder, W.E.; Kreider, B.L.; Gilbert, D.J.; Jenkins,
Mol. Cell. Biol. 14, 4342-4349, 1994
A;Title: Stat4, a novel gamma interferon activation site-binding protein expressed in ea
A;Reference number: A56047; MUID:94277038; PMID:8007943
A;Accession: A56047
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-748 <YAM>
A;Cross-references: UNIPROT:P42228; GB:U09351; NID:g509502; PIDN:AAA19692.1; PID:g509503
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: DNA binding; phosphoprotein

  Query Match             57.3%;  Score 283;  DB 2;  Length 748;
  Best Local Similarity   58.9%;  Pred. No. 6.1e-23;
  Matches   53;  Conservative   15;  Mismatches   22;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              :||:|   |::|:|||   |  | ||   |  || |::|||||||||| || ||:|:||
Db        277 EKLQEQSTKMTYEGDPIPAQRAHLLERATFLIYNLFKNSFVVERQPCMPTHPQRPMVLKT 336

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKD 90
              :||| |:|||:| ||||||:|:|  |||:
Db        337 LIQFTVKLRLLIKLPELNYQVKVKASIDKN 366


RESULT 5
A46160
interferon alpha-induced transcription activator ISGF-3, 113K chain - human
N;Alternate names: stat2 protein
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A46160; S71908; S53873
R;Fu, X.Y.; Schindler, C.; Improta, T.; Aebersold, R.; Darnell Jr., J.E.
Proc. Natl. Acad. Sci. U.S.A. 89, 7840-7843, 1992
A;Title: The proteins of ISGF-3, the interferon alpha-induced transcriptional activator
A;Reference number: A46160; MUID:92366558; PMID:1502204
A;Accession: A46160
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA; protein
A;Residues: 1-851 <FU1>
A;Cross-references: UNIPROT:P52630
A;Note: sequence extracted from NCBI backbone (NCBIP:110820)
R;Yan, R.; Qureshi, S.; Zhong, Z.; Wen, Z.; Darnell, J.E.
submitted to the EMBL Data Library, December 1994
A;Reference number: S71908
A;Accession: S71908
A;Molecule type: DNA
A;Residues: 1-851 <YAN>
A;Cross-references: EMBL:U18671; NID:g1293919; PIDN:AAA98760.1; PID:g1293920
R;Yan, R.; Qureshi, S.; Zhong, Z.; Wen, Z.; Darnell Jr., J.E.
Nucleic Acids Res. 23, 459-463, 1995
A;Title: The genomic structure of the STAT genes: multiple exons in coincident sites in
A;Reference number: S53873; MUID:95192056; PMID:7885841
A;Accession: S53873
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-196;392-591;684-730 <YAW>
A;Cross-references: EMBL:U18671
C;Genetics:
A;Gene: stat2
A;Introns: 44/2; 95/3; 127/3; 157/3; 183/1; 211/3; 261/2; 314/2; 345/2; 365/2; 372/2; 4
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: signal transduction; transcription regulation

  Query Match             41.3%;  Score 204;  DB 2;  Length 851;
  Best Local Similarity   44.4%;  Pred. No. 3.3e-14;
  Matches   40;  Conservative   19;  Mismatches   31;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              |:|: |   |||: ||: :   :   :: || : |:  ||||| ||||| | |||::||
Db        277 KELKGLSCLVSYQDDPLTKGVDLRNAQVTELLQRLLHRAFVVETQPCMPQTPHRPLILKT 336

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKD 90
              | :|| : ||||:  | |  | ::| ||::
Db        337 GSKFTVRTRLLVRLQEGNESLTVEVSIDRN 366


RESULT 6
A54740
interleukin-4-induced transcription factor stat - human
C;Species: Homo sapiens (man)
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004
C;Accession: A54740
R;Hou, J.; Schindler, U.; Henzel, W.J.; Ho, T.C.; Brasseur, M.; McKnight, S.L.
```

Science 265, 1701-1706, 1994
A;Title: An interleukin-4-induced transcription factor: IL-4 stat.
A;Reference number: A54740; MUID:94367369; PMID:8085155
A;Accession: A54740
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-848 <HOU>
A;Cross-references: UNIPROT:P42226
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: DNA binding; transcription regulation

Query Match 18.3%; Score 90.5; DB 2; Length 848;
Best Local Similarity 36.2%; Pred. No. 0.091;
Matches 25; Conservative 12; Mismatches 23; Indels 9; Gaps 2;

```
Qy          5 ELQQKVSYKGDPI-VQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQ 63
              :|||:|   |  :  : |  |  |: |: | |: | |:||:||        | |:||  :
Db        237 QLQQEVGAAGGELEPKTRASLTGRLDEVLRTLVTSCFLVEKQP--------PQVLKTQTK 288

Qy         64 FTTKVRLLV 72
              |   || |:
Db        289 FQAGVRFLL 297
```

RESULT 7
I49274
mammary gland factor - mouse
N;Alternate names: STAT5 protein homolog p80
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I49274; S54773; S54727
R;Liu, X.; Robinson, G.W.; Gouilleux, F.; Groner, B.; Hennighausen, L.
Proc. Natl. Acad. Sci. U.S.A. 92, 8831-8835, 1995
A;Title: Cloning and expression of Stat5 and an additional homologue (Stat5b) involved i
A;Reference number: I49273; MUID:96004632; PMID:7568026
A;Accession: I49274
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-786 <RES>
A;Cross-references: UNIPROT:P42232; UNIPROT:Q9JKM1; EMBL:U21110; NID:g747973; PIDN:AAC52
R;Mui, A.L.F.; Wakao, H.; O'Farrell, A.M.; Harada, N.; Miyajima, A.
EMBO J. 14, 1166-1175, 1995
A;Title: Interleukin-3, granulocyte-macrophage colony stimulating factor and interleukin
A;Reference number: S54772; MUID:95237198; PMID:7720707
A;Accession: S54773
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-432,'E',434-786 <MUI>
A;Cross-references: EMBL:Z48539; NID:g758635; PIDN:CAA88420.1; PID:g758636
R;Azam, M.; Erdjument-Bromage, H.; Kreider, B.L.; Xia, M.; Quelle, F.; Basu, R.; Saris,
EMBO J. 14, 1402-1411, 1995
A;Title: Interleukin-3 signals through multiple isoforms of Stat5.
A;Reference number: S54725; MUID:95246733; PMID:7537213
A;Accession: S54727
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-432,'E',434-786 <AZA>
C;Genetics:
A;Gene: Stat5b
C;Superfamily: human signal transducer and transcription activator STAT5A

Query Match 17.5%; Score 86.5; DB 2; Length 786;
Best Local Similarity 30.6%; Pred. No. 0.23;
Matches 22; Conservative 15; Mismatches 26; Indels 9; Gaps 2;

```
Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              :: | | |::   | |: :   :   | ::   |: | |::|:||        | |:||
Db        294 RRAEHLCQQLPIPG-PVEEMLAEVNATITDIISALVTSTFIIEKQP--------PQVLKT 344

Qy         61 GVQFTTKVRLLV 72
                :|   |||||
Db        345 QTKFAATVRLLV 356
```

RESULT 8
S54772
mammary gland factor - mouse
N;Alternate names: stat5 protein
C;Species: Mus musculus (house mouse)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S54772; I49273
R;Mui, A.L.F.; Wakao, H.; O'Farrell, A.M.; Harada, N.; Miyajima, A.
EMBO J. 14, 1166-1175, 1995
A;Title: Interleukin-3, granulocyte-macrophage colony stimulating factor and interleuki
A;Reference number: S54772; MUID:95237198; PMID:7720707
A;Accession: S54772
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-793 <MUI>
A;Cross-references: UNIPROT:P42230; UNIPROT:Q9JIA0; EMBL:Z48538; NID:g758633; PIDN:CAA8
R;Liu, X.; Robinson, G.W.; Gouilleux, F.; Groner, B.; Hennighausen, L.
Proc. Natl. Acad. Sci. U.S.A. 92, 8831-8835, 1995
A;Title: Cloning and expression of Stat5 and an additional homologue (Stat5b) involved
A;Reference number: I49273; MUID:96004632; PMID:7568026
A;Accession: I49273
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-793 <RES>
A;Cross-references: EMBL:U21103; NID:g747971; PIDN:AAA80590.1; PID:g747972
C;Genetics:
A;Gene: Stat5a
C;Superfamily: human signal transducer and transcription activator STAT5A

Query Match 17.5%; Score 86.5; DB 2; Length 793;
Best Local Similarity 30.6%; Pred. No. 0.23;
Matches 22; Conservative 15; Mismatches 26; Indels 9; Gaps 2;

```
Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              :: | | |::   | |: :   :   | ::   |: | |::|:||        | |:||
Db        294 RRAEHLCQQLPIPG-PVEEMLAEVNATITDIISALVTSTFIIEKQP--------PQVLKT 344

Qy         61 GVQFTTKVRLLV 72
                :|   |||||
Db        345 QTKFAATVRLLV 356
```

RESULT 9
G02317
transcription activator stat5A - human
C;Species: Homo sapiens (man)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 09-Jul-2004
C;Accession: G02317
R;Lin, J.
submitted to the EMBL Data Library, December 1995
A;Reference number: H01043
A;Accession: G02317
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-794 <LIN>
A;Cross-references: UNIPROT:P42229; EMBL:U43185; NID:g1151169; PIDN:AAB06589.1; PID:g11
C;Superfamily: human signal transducer and transcription activator STAT5A

Query Match 17.5%; Score 86.5; DB 2; Length 794;
Best Local Similarity 30.6%; Pred. No. 0.23;
Matches 22; Conservative 15; Mismatches 26; Indels 9; Gaps 2;

```
Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              :: | | |::   | |: :   :   | ::   |: | |::|:||        | |:||
Db        294 RRAEHLCQQLPIPG-PVEEMLAEVNATITDIISALVTSTFIIEKQP--------PQVLKT 344

Qy         61 GVQFTTKVRLLV 72
                :|   |||||
Db        345 QTKFAATVRLLV 356
```

RESULT 10

```
S55527
mammary gland factor - sheep
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S55527; S44353
R;Wakao, H.; Gouilleux, F.; Groner, B.
EMBO J. 14, 854-855, 1995
A;Title: Corrigenda. Mammary gland factor (MGF) is a novel member of the cytokine regula
A;Reference number: S55527; MUID:95188889; PMID:7882987
A;Accession: S55527
A;Molecule type: mRNA
A;Residues: 1-794 <WAK>
A;Cross-references: UNIPROT:P42231; EMBL:X78428; NID:g602354; PIDN:CAA55191.1; PID:g6023
A;Note: this is a revision to the sequence from reference S44353
R;Wakao, H.; Gouilleux, F.; Groner, B.
EMBO J. 13, 2182-2191, 1994
A;Title: Mammary gland factor (MGF) is a novel member of the cytokine regulated transcri
A;Reference number: S44353; MUID:94244619; PMID:7514531
A;Accession: S44353
A;Molecule type: mRNA
A;Residues: 1-716,'RHLHGPGSLPSR',729,'P',731,'ASL' <WAW>
A;Cross-references: EMBL:X78428
A;Note: this sequence has been revised in reference S55527
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             17.5%;  Score 86.5;  DB 2;  Length 794;
  Best Local Similarity   30.6%;  Pred. No. 0.23;
  Matches   22;  Conservative   15;  Mismatches   26;  Indels    9;  Gaps    2;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              :: | | |::   | |: :    :   | ::   |: | |::|:||        | |:||
Db        294 RRAEHLCQQLPIPG-PVEEMLAEVNATITDIISALVTSTFIIEKQP--------PQVLKT 344

Qy         61 GVQFTTKVRLLV 72
                :|   |||||
Db        345 QTKFAATVRLLV 356


RESULT 11
I57557
DNA-Binding Protein and transcription factor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C;Accession: I57557
R;Quelle, F.W.; Shimoda, K.; Thierfelder, W.; Fischer, C.L.; Kim, A.; Ruben, S.M.; Cleve
Mol. Cell. Biol. 15, 3336-3343, 1995
A;Title: Cloning of murine Stat6 and human Stat6, Stat proteins that are tyrosine phosph
A;Reference number: I57557; MUID:95280934; PMID:7760829
A;Accession: I57557
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-837 <RES>
A;Cross-references: UNIPROT:P52633; GB:L47650; NID:g1008876; PIDN:AAA79006.1; PID:g10088
C;Genetics:
A;Gene: STAT6
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             16.9%;  Score 83.5;  DB 2;  Length 837;
  Best Local Similarity   33.3%;  Pred. No. 0.52;
  Matches   23;  Conservative   15;  Mismatches   22;  Indels    9;  Gaps    2;

Qy          5 ELQQKV-SYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQ 63
              :| |:: :   |:   : |  |  |: |: | |: |:|:||:||        | |:||  :
Db        237 QLHQEIGAASGELEPKTRASLISRLDEVLRTLVTSSFLVEKQP--------PQVLKTQTK 288

Qy         64 FTTKVRLLV 72
                |   || |:
Db        289 FQAGVRFLL 297


RESULT 12
S32458
4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27) - human
C;Species: Homo sapiens (man)
C;Date: 28-May-1993 #sequence_revision 02-Jun-1994 #text_change 12-Jul-2004
C;Accession: S32458; S40298; S32153
R;Rueetschi, U.; Dellsen, A.; Sahlin, P.; Stenman, G.; Rymo, L.; Lindstedt, S.
Eur. J. Biochem. 213, 1081-1089, 1993
A;Title: Human 4-hydroxyphenylpyruvate dioxygenase. Primary structure and chromosomal l
A;Reference number: S32458; MUID:93279307; PMID:8504803
A;Accession: S32458
A;Molecule type: mRNA
A;Residues: 1-393 <RU1>
A;Cross-references: UNIPROT:P32754; EMBL:X72389; NID:g288104; PIDN:CAA51082.1; PID:g288
A;Note: the authors did not translate the codon for residue 1
A;Accession: S40298
A;Molecule type: protein
A;Residues: 2-23;41-76;82-108;119-142;151-162;167-179;181-217;221-268;271-300;302-310;3
C;Genetics:
A;Gene: GDB:HPD; PPD
A;Cross-references: GDB:135978; OMIM:276710
A;Map position: 12q24-12qter
C;Superfamily: 4-hydroxyphenylpyruvate_dioxygenase
C;Keywords: homodimer; oxidoreductase

  Query Match             14.6%;  Score 72;  DB 2;  Length 393;
  Best Local Similarity   24.7%;  Pred. No. 4;
  Matches   22;  Conservative   16;  Mismatches   45;  Indels    6;  Gaps    1;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDR------ 54
              ||  ::|: | | |   |||  :  | |:   |:| :        |        |
Db        247 KKKSQIQEYVDYNGGAGVQHIALKTEDIITAIRHLRERGLEFLSVPSTYYKQLREKLKTA 306

Qy         55 PLVIKTGVQFTTKVRLLVKFPELNYQLKI 83
                : :|  :    ::::|| : |  | |:|
Db        307 KIKVKENIDALEELKILVDYDEKGYLLQI 335


RESULT 13
S32821
4-hydroxyphenylpyruvate dioxygenase (EC 1.13.11.27) - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 12-Jul-2004
C;Accession: S32821; S35890; S35889
R;Endo, F.; Awata, H.; Tanoue, A.; Ishiguro, M.; Eda, Y.; Titani, K.; Matsuda, I.
J. Biol. Chem. 267, 24235-24240, 1992
A;Title: Primary structure deduced from complementary DNA sequence and expression in cu
entical subunits homologous to rat liver-specific alloantigen F.
A;Reference number: S32821; MUID:93077502; PMID:1339442
A;Accession: S32821
A;Molecule type: mRNA
A;Residues: 1-393 <END1>
A;Cross-references: UNIPROT:Q02110; EMBL:D13390
A;Accession: S35890
A;Molecule type: protein
A;Residues: 83-118;151-174;194-196;198-224;230-254;256-265;267-273;341-356;358-361;370-
R;Endo, F.; Awata, H.; Tanoue, A.; Ishiguro, M.; Eda, Y.; Titani, K.; Matsuda, I.
submitted to the EMBL Data Library, October 1992
A;Description: Primary structure deduced from complementary DNA sequence and expression
A;Reference number: S35889
A;Accession: S35889
A;Molecule type: mRNA
A;Residues: 1-120,'EVCCAADVRGHHTPLDRARQVWEG',145-150,'TFCLDSR',158,'QPSQTLLHRL',169-393
A;Cross-references: EMBL:D13390; NID:g217691; PIDN:BAA02660.1; PID:g217692
C;Superfamily: 4-hydroxyphenylpyruvate_dioxygenase
C;Keywords: homodimer; oxidoreductase

  Query Match             14.6%;  Score 72;  DB 2;  Length 393;
  Best Local Similarity   24.4%;  Pred. No. 4;
  Matches   22;  Conservative   17;  Mismatches   43;  Indels    8;  Gaps    2;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSA-------FVVERQPCMPMHPD 53
              ||  ::|: | | |   |||  :  | |:   |:| :          |   :|
Db        247 KKKSQIQEYVDYNGGAGVQHIALKTEDIITAIRSLRERGVEFLAVPFTYYKQ-LQEKLKS 305

Qy         54 RPLVIKTGVQFTTKVRLLVKFPELNYQLKI 83
```

```
              : :|  :    ::::|| : |  | |:|
Db        306 AKIRVKESIDVLEELKILVDYDEKGYLLQI 335


RESULT 14
T06667
argininosuccinate synthase (EC 6.3.4.5) - Arabidopsis thaliana
N;Alternate names: protein F6I7.40
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 22-Oct-1999
C;Accession: T06667
R;Bevan, M.; Terryn, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De Clerck, R.; De
oheisel, J.; Mewes, H.W.; Mayer, K.F.X.; Schueller, C.
submitted to the Protein Sequence Database, April 1999
A;Reference number: Z15792
A;Accession: T06667
A;Molecule type: DNA
A;Residues: 1-498 <BEV>
A;Cross-references: EMBL:AL049657; GSPDB:GN00062; ATSP:F6I7.40
A;Experimental source: cultivar Columbia; BAC clone F6I7
C;Genetics:
A;Gene: ATSP:F6I7.40
A;Map position: 4
A;Introns: 25/1; 100/1; 146/2; 162/3; 223/3; 285/3; 315/2
C;Superfamily: argininosuccinate synthase
C;Keywords: ligase

  Query Match             14.4%;  Score 71;  DB 2;  Length 498;
  Best Local Similarity   25.3%;  Pred. No. 6.7;
  Matches   21;  Conservative   18;  Mismatches   36;  Indels    8;  Gaps    1;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              |:|| |:||    |   :  : : || : :     :::  : ||:  :     ||::
Db        165 KELEGLEQKAKASGASQLVVKDLTEEFVKDFIFPCLRAGAIYERKYLLGTSMARPVI--- 221

Qy         61 GVQFTTKVRLLVKFPELNYQLKI 83
                  |||  : |  || :||:
Db        222 -----AKVRFELTFFSLNPELKV 239


RESULT 15
S48483
MRS1 protein - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YIR021w
C;Species: Saccharomyces cerevisiae
C;Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 09-Jul-2004
C;Accession: S48483; A29684
R;Rowley, K.
submitted to the EMBL Data Library, October 1994
A;Reference number: S48478
A;Accession: S48483
A;Molecule type: DNA
A;Residues: 1-363 <ROW>
A;Cross-references: UNIPROT:P07266; GB:Z47047; EMBL:Z38061; NID:g603997; PID:g763366; MI
R;Kreike, J.; Schulze, M.; Ahne, F.; Lang, B.F.
EMBO J. 6, 2123-2129, 1987
A;Title: A yeast nuclear gene, MRS1, involved in mitochondrial RNA splicing: nucleotide
A;Reference number: A91075; MUID:88004424; PMID:2443348
A;Accession: A29684
A;Molecule type: DNA
A;Residues: 1-181,'P',183-317,'A',319-363 <KRE>
A;Cross-references: EMBL:X05509; NID:g3987; PIDN:CAA29053.1; PID:g3988
C;Genetics:
A;Gene: SGD:MRS1; PET157
A;Cross-references: SGD:S0001460; MIPS:YIR021w
A;Map position: 9R
A;Genome: nuclear
C;Keywords: mitochondrion

  Query Match             13.8%;  Score 68;  DB 2;  Length 363;
  Best Local Similarity   26.0%;  Pred. No. 10;
  Matches   25;  Conservative   21;  Mismatches   32;  Indels   18;  Gaps    4;

Qy          9 KVSYKGDPIV-------QHRPMLEERIVELF-RNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              :|| | :| :         : |::    | |:  : |::|||  :  | | :| |:   :|
Db        210 RVSLKENPSISVPREYQMYDPVVRATIKEVVTKRLLRSAFDNDIDPLMCLHLDKGWKLKF 269

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCID--KDSGDV 94
                :  :|          ||: ||  : :|  ||: |:
Db        270 PILSST--------TGLNFSLKDCLSLDTGKDASDM 297


Search completed: May 25, 2005, 17:45:27
Job time : 11.0566 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:15:30 ; Search time 39.9444 Seconds
(without alignments)
1230.701 Million cell updates/sec

Title: US-10-090-185-25
Perfect score: 494
Sequence: 1 KKLEELQQKVSYKGDPIVQH.....LNYQLKIKVCIDKDSGDVAA 96

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	494	100.0	163	2	Q9N145	Q9n145 macaca mula
2	494	100.0	770	1	STA3_BOVIN	P61635 bos taurus
3	494	100.0	770	1	STA3_HUMAN	P40763 homo sapien
4	494	100.0	770	1	STA3_MOUSE	P42227 mus musculu
5	491	99.4	770	1	STA3_RAT	P52631 rattus norv
6	490	99.2	771	2	Q6DV79	Q6dv79 gallus gall
7	489	99.0	769	2	Q9PVX8	Q9pvx8 xenopus lae
8	487	98.6	161	2	Q704W5	Q704w5 bos taurus
9	475	96.2	766	2	Q7ZXK3	Q7zxk3 xenopus lae
10	474	96.0	414	2	Q7ZTS5	Q7zts5 brachydanio
11	474	96.0	786	2	Q6NV46	Q6nv46 brachydanio
12	474	96.0	806	2	O93599	O93599 brachydanio
13	472	95.5	765	2	Q6DVF3	Q6dvf3 oryzias lat
14	472	95.5	785	2	Q6GUE7	Q6gue7 oryzias lat
15	461	93.3	764	2	Q90Y16	Q90y16 tetraodon f
16	453	91.7	767	2	O13133	O13133 oncorhynchu
17	315	63.8	751	2	Q8JGN0	Q8jgn0 xenopus lae
18	311	63.0	712	2	Q6P6Q7	Q6p6q7 rattus norv
19	311	63.0	749	2	Q8C497	Q8c497 mus musculu
20	311	63.0	749	2	Q9QXK0	Q9qxk0 rattus norv
21	310	62.8	712	2	Q99K94	Q99k94 mus musculu
22	310	62.8	749	2	Q8C3V4	Q8c3v4 mus musculu
23	310	62.8	749	2	Q9D323	Q9d323 mus musculu
24	310	62.8	750	1	STA1_HUMAN	P42224 homo sapien
25	310	62.8	750	2	Q68D00	Q68d00 homo sapien
26	310	62.8	755	2	Q8C8M3	Q8c8m3 mus musculu
27	309	62.6	758	2	Q90Y17	Q90y17 tetraodon f
28	304	61.5	757	2	Q764M5	Q764m5 sus scrofa
29	294	59.5	1153	2	Q8JFS5	Q8jfs5 brachydanio
30	292	59.1	754	2	O13131	O13131 oncorhynchu
31	292	59.1	754	2	O13132	O13132 oncorhynchu
32	291	58.9	748	1	STA4_HUMAN	Q14765 homo sapien
33	290	58.7	718	2	Q801Y2	Q801y2 carassius a
34	288	58.3	528	2	Q8JFU8	Q8jfu8 brachydanio
35	285	57.7	737	2	Q90Y15	Q90y15 tetraodon f
36	283	57.3	749	1	STA4_MOUSE	P42228 mus musculu
37	282	57.1	748	2	Q66HB2	Q66hb2 rattus norv
38	280	56.7	553	2	Q8JFU7	Q8jfu7 brachydanio
39	280	56.7	651	2	Q7ZZ53	Q7zz53 brachydanio
40	280	56.7	652	2	Q7ZZ77	Q7zz77 brachydanio
41	280	56.7	657	2	Q8AW24	Q8aw24 brachydanio
42	280	56.7	667	2	Q8AW20	Q8aw20 brachydanio
43	279	56.5	749	1	STA1_MOUSE	P42225 mus musculu
44	261	52.8	749	2	O93598	O93598 brachydanio
45	261	52.8	749	2	Q6P943	Q6p943 brachydanio

ALIGNMENTS

RESULT 1
Q9N145
ID Q9N145 PRELIMINARY; PRT; 163 AA.
AC Q9N145;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Signal transducer and activator of transcription (Fragment).
GN Name=STAT3;
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Arredondo J.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF227560; AAF73401.1; -.
DR HSSP; P42227; 1BG1.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR008967; P53_like_DNA_bnd.
DR InterPro; IPR001217; STAT.
DR Pfam; PF01017; STAT_alpha; 1.
DR Pfam; PF02864; STAT_bind; 1.
FT NON_TER 1 1
FT NON_TER 163 163
SQ SEQUENCE 163 AA; 18855 MW; DD3F11E7F5B9878C CRC64;

Query Match 100.0%; Score 494; DB 2; Length 163;
Best Local Similarity 100.0%; Pred. No. 1.3e-46;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         52 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 111

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        112 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 147
```

RESULT 2
STA3_BOVIN
ID STA3_BOVIN STANDARD; PRT; 770 AA.
AC P61635;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Signal transducer and activator of transcription 3.

```
GN   Name=STAT3;
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Mammary gland;
RA   Seyfert H.M., Wheeler T.T., Moolenaar A., Pitra C.;
RT   "The STAT5B-encoding gene was flipped across the STAT3/STAT5A-locus
RT   during ruminant evolution.";
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes (By similarity).
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation (By similarity).
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ620655; CAF06182.1; -.
DR   PROSITE; PS50001; SH2; 1.
KW   Activator; DNA-binding; Nuclear protein; Phosphorylation; SH2 domain;
KW   Transcription regulation.
FT   DOMAIN      580    670       SH2.
FT   MOD_RES     705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES     727    727       Phosphoserine (By similarity).
SQ   SEQUENCE   770 AA;  87974 MW;  9CEB147C73E83274 CRC64;

  Query Match             100.0%;  Score 494;  DB 1;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 7.7e-46;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 3
STA3_HUMAN
ID   STA3_HUMAN     STANDARD;      PRT;   770 AA.
AC   P40763; O14916; Q9BW54;
DT   01-FEB-1995 (Rel. 31, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Signal transducer and activator of transcription 3 (Acute-phase
DE   response factor).
GN   Name=STAT3; Synonyms=APRF;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Placenta;
RX   MEDLINE=94208062; PubMed=7512451; DOI=10.1016/0092-8674(94)90235-6;
RA   Akira S., Nishio Y., Inoue M., Wang X.-J., Wei S., Matsusaka T.,
RA   Yoshida K., Sudo T., Naruto M., Kishimoto T.;
RT   "Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-
RT   related transcription factor involved in the gp130-mediated signaling
RT   pathway.";
RL   Cell 77:63-71(1994).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=98296260; PubMed=9630560; DOI=10.1016/S0378-1119(98)00185-1;
RA   Della Pietra L., Bressan A., Pezzotti A., Serlupi-Crescenzi O.;
RT   "Highly conserved amino-acid sequence between murine STAT3 and a
RT   revised human STAT3 sequence.";
RL   Gene 213:119-124(1998).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM DEL-701), AND VARIANT ILE-143.
RA   Rieder M.J., Daniels R.L., da Ponte S.H., Hastings N.C., Ahearn M.O.,
RA   Rajkumar N., Yi Q., Nickerson D.A.;
RT   "SeattleSNPs. NHLBI HL66682 program for genomic applications, UW-
RT   FHCRC, Seattle, WA (URL: http://pga.gs.washington.edu).";
RL   Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND DEL-701).
RC   TISSUE=Kidney, and Pancreas;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   SEQUENCE OF 564-704 FROM N.A.
RC   TISSUE=Liver;
RA   Della Pietra L., Bressan A., Pezzotti A.R., Serlupi-Crescenzi O.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   PHOSPHORYLATION ON SERINE.
RX   MEDLINE=95215843; PubMed=7701321;
RA   Zhang X., Blenis J., Li H.-C., Schindler C., Chen-Kiang S.;
RT   "Requirement of serine phosphorylation for formation of STAT-promoter
RT   complexes.";
RL   Science 267:1990-1994(1995).
RN   [7]
RP   INTERACTION WITH NCOA1.
RX   PubMed=11773079; DOI=10.1074/jbc.M111486200;
RA   Giraud S., Bienvenu F., Avril S., Gascan H., Heery D.M., Coqueret O.;
RT   "Functional interaction of STAT3 transcription factor with the
RT   coactivator NcoA/SRC1a.";
RL   J. Biol. Chem. 277:8004-8011(2002).
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes.
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
```

```
CC         IsoId=P40763-1; Sequence=Displayed;
CC       Name=Del-701;
CC         IsoId=P40763-2; Sequence=VSP_010474;
CC   -!- TISSUE SPECIFICITY: Heart, brain, placenta, lung, liver, skeletal
CC       muscle, kidney and pancreas.
CC   -!- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,
CC       LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation
CC       is important for the formation of stable DNA-binding STAT3
CC       homodimers and maximal transcriptional activity.
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L29277; AAA58374.1; -.
DR   EMBL; AJ012463; CAA10032.1; -.
DR   EMBL; AY572796; AAS66986.1; -.
DR   EMBL; BC000627; AAH00627.1; -.
DR   EMBL; BC014482; AAH14482.1; -.
DR   EMBL; AF029311; AAB84254.1; -.
DR   PIR; A54444; A54444.
DR   HSSP; P42227; 1BG1.
DR   TRANSFAC; T01493; -.
DR   Genew; HGNC:11364; STAT3.
DR   H-InvDB; HIX0013840; -.
DR   MIM; 102582; -.
DR   GO; GO:0005737; C:cytoplasm; TAS.
DR   GO; GO:0005634; C:nucleus; TAS.
DR   GO; GO:0005062; F:hematopoietin/interferon-class (D200-domain. . .; TAS.
DR   GO; GO:0003700; F:transcription factor activity; TAS.
DR   GO; GO:0006928; P:cell motility; TAS.
DR   GO; GO:0007259; P:JAK-STAT cascade; TAS.
DR   GO; GO:0000122; P:negative regulation of transcription from P. . .; TAS.
DR   GO; GO:0007399; P:neurogenesis; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
KW   Activator; Alternative splicing; DNA-binding; Nuclear protein;
KW   Phosphorylation; Polymorphism; SH2 domain; Transcription regulation.
FT   DOMAIN      580    670       SH2.
FT   MOD_RES     705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES     727    727       Phosphoserine (By similarity).
FT   VARSPLIC    701    701       Missing (in isoform Del-701).
FT                                /FTId=VSP_010474.
FT   VARIANT      32     32       Q -> K (in dbSNP:1803125).
FT                                /FTId=VAR_018683.
FT   VARIANT     143    143       M -> I.
FT                                /FTId=VAR_018679.
FT   CONFLICT    288    288       Q -> H (in Ref. 1).
FT   CONFLICT    460    460       P -> S (in Ref. 1).
FT   CONFLICT    548    548       K -> N (in Ref. 1).
FT   CONFLICT    561    561       F -> Y (in Ref. 1).
FT   CONFLICT    667    667       V -> L (in Ref. 1).
FT   CONFLICT    730    730       T -> A (in Ref. 1).
SQ   SEQUENCE   770 AA;  88067 MW;  6C00632211C8012D CRC64;

  Query Match             100.0%;  Score 494;  DB 1;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 7.7e-46;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 4
STA3_MOUSE
ID   STA3_MOUSE     STANDARD;      PRT;   770 AA.
AC   P42227;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Signal transducer and activator of transcription 3 (Acute-phase
DE   response factor).
GN   Name=Stat3; Synonyms=Aprf;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A), AND SEQUENCE OF 154-158; 181-185
RP   AND 632-640.
RC   STRAIN=BALB/c; TISSUE=Liver;
RX   MEDLINE=94208062; PubMed=7512451; DOI=10.1016/0092-8674(94)90235-6;
RA   Akira S., Nishio Y., Inoue M., Wang X.J., Shi W., Matsusaka T.,
RA   Yoshida K., Sudo T., Naruto M., Kishimoto T.;
RT   "Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-
RT   related transcription factor involved in the gp130-mediated signaling
RT   pathway.";
RL   Cell 77:63-71(1994).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   TISSUE=Thymus;
RX   MEDLINE=94188718; PubMed=8140422;
RA   Zhong Z., Wen Z., Darnell J.E. Jr.;
RT   "Stat3: a STAT family member activated by tyrosine phosphorylation in
RT   response to epidermal growth factor and interleukin-6.";
RL   Science 264:95-98(1994).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   TISSUE=Brain;
RX   MEDLINE=95014185; PubMed=7523373;
RA   Raz R., Durbin J.E., Levy D.E.;
RT   "Acute phase response factor and additional members of the interferon-
RT   stimulated gene factor 3 family integrate diverse signals from
RT   cytokines, interferons, and growth factors.";
RL   J. Biol. Chem. 269:24391-24395(1994).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3B).
RC   STRAIN=BALB/c, and C57BL/6; TISSUE=Liver;
RX   MEDLINE=96016116; PubMed=7568080;
RA   Schaefer T.S., Sanders L.K., Nathans D.;
RT   "Cooperative transcriptional activity of Jun and Stat3 beta, a short
RT   form of Stat3.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:9097-9101(1995).
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   STRAIN=129/SvJ;
RX   PubMed=11161808; DOI=10.1006/geno.2000.6433;
RA   Miyoshi K., Cui Y., Riedlinger G., Lehoczky J., Zon L., Oka T.,
RA   Dewar K., Hennighausen L.;
RT   "Structure of the mouse stat 3/5 locus: evolution from Drosophila to
RT   zebrafish to mouse.";
RL   Genomics 71:150-155(2001).
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   STRAIN=C57BL/6J, and NOD/LtJ;
RA   Davoodi-Semiromi A., She J.-X.;
RT   "A mutant Stat5b with weaker DNA binding defines a key defective
RT   pathway in non-obese diabetic (NOD) mice.";
```

```
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   STRAIN=FVB/N; TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [8]
RP   PHOSPHORYLATION SITE SER-727, AND MUTAGENESIS.
RX   MEDLINE=95354205; PubMed=7543024; DOI=10.1016/0092-8674(95)90311-9;
RA   Wen Z., Zhong Z., Darnell J.E. Jr.;
RT   "Maximal activation of transcription by Stat1 and Stat3 requires both
RT   tyrosine and serine phosphorylation.";
RL   Cell 82:241-250(1995).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 136-716.
RX   MEDLINE=98334373; PubMed=9671298; DOI=10.1038/28101;
RA   Becker S., Groner B., Mueller C.W.;
RT   "Three-dimensional structure of the Stat3beta homodimer bound to
RT   DNA.";
RL   Nature 394:145-151(1998).
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes. STAT3B interacts with the N-terminal
CC       part of JUN to activate such promoters in a cooperative way.
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=Stat3A;
CC         IsoId=P42227-1; Sequence=Displayed;
CC       Name=Stat3B;
CC         IsoId=P42227-2; Sequence=VSP_006287;
CC       Name=Del-701;
CC         IsoId=P42227-3; Sequence=VSP_010475;
CC   -!- TISSUE SPECIFICITY: STAT3A is seen in the liver, spleen, and
CC       kidney. STAT3B is also detected in the liver, although in a much
CC       less abundant manner.
CC   -!- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,
CC       LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation
CC       is important for the formation of stable DNA-binding STAT3
CC       homodimers and maximal transcriptional activity (By similarity).
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L29278; AAA37254.1; -.
DR   EMBL; U06922; AAA19452.1; -.
DR   EMBL; U08378; AAA56668.1; -.
DR   EMBL; U30709; AAC52612.1; -.
DR   EMBL; AF246978; AAL59017.1; -.
DR   EMBL; AY299489; AAQ75418.1; -.
DR   EMBL; AY299490; AAQ75419.1; -.
DR   EMBL; BC003806; AAH03806.1; -.
DR   PIR; I49508; I49508.
DR   PDB; 1BG1; X-ray; A=1-722.
DR   TRANSFAC; T01574; -.
DR   MGD; MGI:103038; Stat3.
DR   GO; GO:0005737; C:cytoplasm; IDA.
DR   GO; GO:0005634; C:nucleus; IDA.
DR   GO; GO:0005886; C:plasma membrane; IDA.
DR   GO; GO:0003677; F:DNA binding; IDA.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   GO; GO:0016563; F:transcriptional activator activity; IDA.
DR   GO; GO:0007259; P:JAK-STAT cascade; IDA.
DR   GO; GO:0006357; P:regulation of transcription from Pol II pro. . .; IDA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
KW   3D-structure; Activator; Acute phase; Alternative splicing;
KW   Direct protein sequencing; DNA-binding; Nuclear protein;
KW   Phosphorylation; SH2 domain; Transcription regulation.
FT   DOMAIN      580    670       SH2.
FT   MOD_RES     705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES     727    727       Phosphoserine.
FT   VARSPLIC    716    770       TTCSNTIDLPMSPRTLDSLMQFGNNGEGAEPSAGGQFESLT
FT                                FDMDLTSECATSPM -> FIDAVWK (in isoform
FT                                Stat3B).
FT                                /FTId=VSP_006287.
FT   VARSPLIC    701    701       Missing (in isoform Del-701).
FT                                /FTId=VSP_010475.
FT   MUTAGEN     727    727       S->A: Decreased transcriptional
FT                                activation.
FT   CONFLICT     16     16       E -> K (in Ref. 2).
FT   CONFLICT     25     25       S -> T (in Ref. 2 and 4).
FT   CONFLICT    394    394       M -> I (in Ref. 1).
FT   HELIX       139    180
FT   TURN        181    182
FT   TURN        197    198
FT   HELIX       199    237
FT   TURN        238    238
FT   HELIX       239    251
FT   TURN        252    253
FT   HELIX       261    290
FT   TURN        294    295
FT   TURN        297    301
FT   HELIX       302    320
FT   STRAND      321    328
FT   TURN        330    331
FT   TURN        333    334
FT   TURN        336    337
FT   STRAND      338    340
FT   TURN        341    342
FT   STRAND      345    351

  Query Match          100.0%;  Score 494;  DB 1;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 7.7e-46;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341
```

```
Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              |||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 5
STA3_RAT
ID   STA3_RAT       STANDARD;      PRT;   770 AA.
AC   P52631;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Signal transducer and activator of transcription 3.
GN   Name=Stat3;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=96102059; PubMed=8530402; DOI=10.1074/jbc.270.50.29998;
RA   Ripperger J.A., Fritz S., Richter K., Hocke G.M., Lottspeich F.,
RA   Fey G.H.;
RT   "Transcription factors Stat3 and Stat5b are present in rat liver
RT   nuclei late in an acute phase response and bind interleukin-6 response
RT   elements.";
RL   J. Biol. Chem. 270:29998-30006(1995).
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes.
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation (By similarity).
CC   -!- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,
CC       LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation
CC       is important for the formation of stable DNA-binding STAT3
CC       homodimers and maximal transcriptional activity (By similarity).
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X91810; CAA62920.1; -.
DR   HSSP; P42227; 1BG1.
DR   RGD; 3772; Stat3.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
KW   Activator; DNA-binding; Nuclear protein; Phosphorylation; SH2 domain;
KW   Transcription regulation.
FT   DOMAIN      580    670       SH2.
FT   MOD_RES     705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES     727    727       Phosphoserine (By similarity).
SQ   SEQUENCE   770 AA;  88039 MW;  D74A0C76954754ED CRC64;

  Query Match             99.4%;  Score 491;  DB 1;  Length 770;
  Best Local Similarity   99.0%;  Pred. No. 1.7e-45;
  Matches   95;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              |||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 6
Q6DV79
ID   Q6DV79      PRELIMINARY;      PRT;   771 AA.
AC   Q6DV79;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Signal transducer and activator of transcription 3.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zhou G.Y., Leung F.C.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY641397; AAT64887.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   771 AA;  88174 MW;  71AC855C5DEC03E2 CRC64;

  Query Match             99.2%;  Score 490;  DB 2;  Length 771;
  Best Local Similarity   99.0%;  Pred. No. 2.1e-45;
  Matches   95;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVFKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              |||||||||||||||||||||||||||||||||||||
Db        342 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 7
Q9PVX8
ID   Q9PVX8      PRELIMINARY;      PRT;   769 AA.
AC   Q9PVX8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat 3.
GN   Name=stat 3;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20107399; PubMed=10642787; DOI=10.1006/dbio.1999.9518;
```

```
RA   Nishinakamura R., Matsumoto Y., Matsuda T., Ariizumi T., Heike T.,.
RA   Asashima M., Yokota T.;
RT   "Activation of Stat3 by cytokine receptor gp130 ventralizes Xenopus
RT   embryos independent of BMP-4.";
RL   Dev. Biol. 216:481-490(1999).
DR   EMBL; AB017701; BAA86061.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   769 AA;  87974 MW;  0905C03263303069 CRC64;

  Query Match             99.0%;  Score 489;  DB 2;  Length 769;
  Best Local Similarity   99.0%;  Pred. No. 2.8e-45;
  Matches   95;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||| |||||||||||||||||||||||||||||
Db        342 GVQFTNKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 377


RESULT 8
Q704W5
ID   Q704W5      PRELIMINARY;      PRT;   161 AA.
AC   Q704W5;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Signal transducer and activator of transcription 3 (Fragment).
GN   Name=stat3;
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Mammary gland;
RA   Molenaar A., Wheeler T.T., McCracken J.Y., Seyfert H.M.;
RT   "The STAT3-encoding gene resides within the 40 kbp gap between the
RT   STAT5A- and STAT5B-encoding gene in cattle.";
RL   Anim. Genet. 31:333-334(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Mammary gland;
RA   Seyfert H.M., Wheeler T.T., Moolenaar A., Pitra C.;
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ620661; CAF06188.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
FT   NON_TER       1       1
FT   NON_TER     161     161
SQ   SEQUENCE   161 AA;  18342 MW;  5DCEE8F2C3A48191 CRC64;

  Query Match             98.6%;  Score 487;  DB 2;  Length 161;
  Best Local Similarity   99.0%;  Pred. No. 7.7e-46;
  Matches   95;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         16 KKLEGLQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 75

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||||||||||||||||||||||||||||||||||
Db         76 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 111


RESULT 9
Q7ZXK3
ID   Q7ZXK3      PRELIMINARY;      PRT;   766 AA.
AC   Q7ZXK3;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat3-A protein.
GN   Name=stat3-A;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC044717; AAH44717.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
```

```
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   766 AA;  87599 MW;  31018A3321CCEB9C CRC64;

  Query Match             96.2%;  Score 475;  DB 2;  Length 766;
  Best Local Similarity   95.8%;  Pred. No. 9.7e-44;
  Matches   92;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        282 RKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 341

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||| |||||||||||||||||||||: ||
Db        342 GVQFTNKVRLLVKFPELNYQLKIKVCIDKDSGEGAA 377


RESULT 10
Q7ZTS5
ID   Q7ZTS5      PRELIMINARY;      PRT;   414 AA.
AC   Q7ZTS5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat3 protein.
GN   Name=stat3;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB; TISSUE=Whole body;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB; TISSUE=Whole body;
RA   Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC045276; AAH45276.1; -.
DR   HSSP; P42227; 1BG1.
DR   ZFIN; ZDB-GENE-980526-68; stat3.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
SQ   SEQUENCE   414 AA;  48253 MW;  0FFD1B509B7526BD CRC64;

  Query Match             96.0%;  Score 474;  DB 2;  Length 414;
  Best Local Similarity   93.8%;  Pred. No. 6.2e-44;
  Matches   90;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              :|||||||||||||||:|||| |||:||:|||||||||||||||||||||||||||||||
Db        283 RKLEELQQKVSYKGDPIIQHRPALEEKIVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKT 342

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              |||||||||||||||||||||||||||||:||||||
Db        343 GVQFTTKVRLLVKFPELNYQLKIKVCIDKESGDVAA 378


RESULT 11
Q6NV46
ID   Q6NV46      PRELIMINARY;      PRT;   786 AA.
AC   Q6NV46;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Stat3 protein.
GN   Name=stat3;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC068320; AAH68320.1; -.
DR   ZFIN; ZDB-GENE-980526-68; stat3.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
```

```
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   786 AA;  90039 MW;  FC7371D0B0E5447E CRC64;

  Query Match          96.0%;  Score 474;  DB 2;  Length 786;
  Best Local Similarity   93.8%;  Pred. No. 1.3e-43;
  Matches   90;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              :|||||||||||||||||:|||| |||:||:||||||||||||||||||||||||||||||
Db        283 RKLEELQQKVSYKGDPIIQHRPALEEKIVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKT 342

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              |||||||||||||||||||||||||||:||||||
Db        343 GVQFTTKVRLLVKFPELNYQLKIKVCIDKESGDVAA 378


RESULT 12
O93599
ID   O93599      PRELIMINARY;      PRT;   806 AA.
AC   O93599;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Transcription factor.
GN   Name=stat3;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Oates A.C.;
RL   Thesis (1998), University of Melbourne, Australia.
DR   EMBL; AJ005693; CAA06677.1; -.
DR   HSSP; P42227; 1BG1.
DR   ZFIN; ZDB-GENE-980526-68; stat3.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   806 AA;  92151 MW;  74BC4EA401C3C942 CRC64;

  Query Match          96.0%;  Score 474;  DB 2;  Length 806;
  Best Local Similarity   93.8%;  Pred. No. 1.3e-43;
  Matches   90;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              :|||||||||||||||||:|||| |||:||:||||||||||||||||||||||||||||||
Db        283 RKLEELQQKVSYKGDPIIQHRPALEEKIVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKT 342

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              |||||||||||||||||||||||||||:||||||
Db        343 GVQFTTKVRLLVKFPELNYQLKIKVCIDKESGDVAA 378


RESULT 13
Q6DVF3
ID   Q6DVF3      PRELIMINARY;      PRT;   765 AA.
AC   Q6DVF3;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Signal transducer and activation of transcription factor 3.
OS   Oryzias latipes (Medaka fish) (Japanese ricefish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Liu R., Hong Y.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY639947; AAT64912.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   765 AA;  87566 MW;  F5D01408748EC703 CRC64;

  Query Match          95.5%;  Score 472;  DB 2;  Length 765;
  Best Local Similarity   93.8%;  Pred. No. 2.1e-43;
  Matches   90;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              |||||||||||||||||:|||| |||:||:||||||||||||||||||||||||||||||
Db        283 KKLEELQQKVSYKGDPIIQHRPALEEKIVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKT 342

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||| |||||||||||||||||||:||||||
Db        343 GVQFTNKVRLLVKFPELNYQLKIKVCIDKESGDVAA 378


RESULT 14
Q6GUE7
ID   Q6GUE7      PRELIMINARY;      PRT;   785 AA.
AC   Q6GUE7;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Signal transducer and activator of transcription 3 isoform 1.
OS   Oryzias latipes (Medaka fish) (Japanese ricefish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Liu R., Hong Y.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY641434; AAT46364.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
```

```
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   785 AA;  89643 MW;  81F231BDE27DE938 CRC64;

  Query Match             95.5%;  Score 472;  DB 2;  Length 785;
  Best Local Similarity   93.8%;  Pred. No. 2.2e-43;
  Matches   90;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              |||||||||||||||||:|||| ||| :|| :|||||||||||||||||||||||||||||
Db        283 KKLEELQQKVSYKGDPIIQHRPALEEKIVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKT 342

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||| |||||||||||||||||||||:||||||
Db        343 GVQFTNKVRLLVKFPELNYQLKIKVCIDKESGDVAA 378


RESULT 15
Q90Y16
ID   Q90Y16      PRELIMINARY;      PRT;   764 AA.
AC   Q90Y16;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   STAT3.
OS   Tetraodon fluviatilis (Puffer fish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=47145;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sung S.C., Fan T.J., Leu J.H., Chou C.M., Huang C.J.;
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF307106; AAL09415.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   764 AA;  87411 MW;  E661FFE18BEFD8BE CRC64;

  Query Match             93.3%;  Score 461;  DB 2;  Length 764;
  Best Local Similarity   91.7%;  Pred. No. 3.5e-42;
  Matches   88;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 KKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKT 60
              |||||||||||||||||:|||| ||| :|| :|||||||||||||||||||||||||||||
Db        283 KKLEELQQKVSYKGDPIIQHRPALEEKIVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKT 342

Qy         61 GVQFTTKVRLLVKFPELNYQLKIKVCIDKDSGDVAA 96
              ||||| |||||||| |||||||||||||||:|||||:
Db        343 GVQFTNKVRLLVKFQELNYQLKIKVCIDKESGDVAS 378


Search completed: May 25, 2005, 17:43:42
Job time : 40.9444 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:14:45 ; Search time 121.891 Seconds
(without alignments)
726.619 Million cell updates/sec

Title: US-10-090-185-29
Perfect score: 1169
Sequence: 1 NHPTAAVVTEKQQMLEQHLQ.....IKTGVQFATAVALLVKFPEL 229

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1160	99.2	229	4	AAY72860	Aay72860 Mouse Sta
2	1155	98.8	229	4	AAY72850	Aay72850 Mouse Sta
3	1155	98.8	252	4	AAY72846	Aay72846 Mouse Sta
4	1155	98.8	271	4	AAY72841	Aay72841 Mouse Sta
5	1155	98.8	770	2	AAR72082	Aar72082 Mouse Sta
6	1155	98.8	770	2	AAW03176	Aaw03176 Mouse STA
7	1151	98.5	229	4	AAY72863	Aay72863 Mouse Sta
8	1150	98.4	229	4	AAY72862	Aay72862 Mouse Sta
9	1144	97.9	720	5	AAE22055	Aae22055 Human Sta
10	1144	97.9	769	5	ABB57164	Abb57164 Mouse isc
11	1144	97.9	769	5	AAE22054	Aae22054 Human Sta
12	1144	97.9	769	5	AAE22056	Aae22056 Human pro
13	1144	97.9	770	2	AAR82995	Aar82995 Mouse liv
14	1144	97.9	770	2	AAY03768	Aay03768 Human STA
15	1144	97.9	770	3	AAB12377	Aab12377 N-termina
16	1144	97.9	770	5	AAE14652	Aae14652 Murine ST
17	1144	97.9	770	5	ABG69497	Abg69497 Human bai
18	1144	97.9	770	6	ABU10476	Abu10476 Mouse STA
19	1144	97.9	770	8	ADN04365	Adn04365 Antipsori
20	1144	97.9	770	8	ADP54789	Adp54789 Human PRO
21	1144	97.9	793	3	AAB58442	Aab58442 Lung canc
22	1141	97.6	770	7	ADD44738	Add44738 Rat Prote
23	1140.5	97.6	228	4	AAY72861	Aay72861 Mouse Sta
24	1139	97.4	770	2	AAR82993	Aar82993 Human pla
25	1139	97.4	770	4	AAB19964	Aab19964 Human sig
26	1139	97.4	770	5	AAE15174	Aae15174 Human Sta
27	1139	97.4	770	7	ADD44740	Add44740 Human Pro
28	1094	93.6	213	4	AAY72851	Aay72851 Mouse Sta
29	1094	93.6	236	4	AAY72847	Aay72847 Mouse Sta
30	1027	87.9	223	4	AAY72854	Aay72854 Mouse Sta
31	828	70.8	185	4	AAY72855	Aay72855 Mouse Sta
32	781	66.8	176	4	AAY72848	Aay72848 Mouse Sta
33	653	55.9	128	4	AAY72852	Aay72852 Mouse Sta
34	605	51.8	143	4	AAY72849	Aay72849 Mouse Sta
35	560	47.9	749	5	AAG78526	Aag78526 Rat STAT-
36	556	47.6	268	4	AAY72844	Aay72844 Mouse Sta
37	556	47.6	582	2	AAW62996	Aaw62996 Human tru
38	556	47.6	582	6	ABU04748	Abu04748 Human exp
39	556	47.6	582	8	ADH57036	Adh57036 Truncated
40	556	47.6	680	6	ABR59713	Abr59713 Human sig
41	556	47.6	712	2	AAR72079	Aar72079 Human Sta
42	556	47.6	712	2	AAW03170	Aaw03170 Human STA
43	556	47.6	712	2	AAW62995	Aaw62995 Human Sta
44	556	47.6	712	6	ABU04747	Abu04747 Human exp
45	556	47.6	712	6	ABU04735	Abu04735 Human exp

ALIGNMENTS

RESULT 1
AAY72860
ID AAY72860 standard; protein; 229 AA.
XX
AC AAY72860;
XX
DT 31-MAY-2001 (first entry)
XX
DE Mouse Stat3 mutant (L148A,V151A,T346A,K348A,R350A) protein fragment.
XX
KW Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW cellular transformation; dysproliferative disease; cancer; psoriasis;
KW therapy; mutant; mutein.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT Region 1. .25
FT /note= "Stat3-c-Jun interaction region 1; corresponds to
FT 130-154 position of Stat3 protein"
FT Misc-difference 19
FT /note= "Wild type Leu substituted with Ala; corresponds
FT to 148 position of Stat-3 protein"
FT Misc-difference 22
FT /note= "Wild type Val substituted with Ala corresponds to
FT 151 position of Stat-3 protein"
FT Region 213. .229
FT /note= "Stat3-c-Jun interaction region 2; corresponds to
FT 342-358 position of Stat3 protein"
FT Misc-difference 217
FT /note= "Wild type Thr substituted with Ala corresponds to
FT 346 position of Stat-3 protein"
FT Misc-difference 219
FT /note= "Wild type Lys substituted with Ala corresponds to
FT 348 position of Stat-3 protein"
FT Misc-difference 221
FT /note= "Wild type Arg substituted with Ala corresponds to
FT 350 position of Stat-3 protein"
XX
PN WO200116605-A2.
XX
PD 08-MAR-2001.
XX
PF 30-AUG-2000; 2000WO-US023822.
XX
PR 31-AUG-1999; 99US-00387418.
XX
PA (UYRQ) UNIV ROCKEFELLER.

```
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Example 4; Page; 86pp; English.
XX
CC   The present sequence is mouse Stat3 mutant protein fragment containing
CC   130-358 amino acids of Stat3 protein. This mutant is obtained by
CC   replacing Leu 148 with Ala, Val 151 with Ala, Thr 346 with Ala, Lys 348
CC   with Ala and Arg 350 with Ala in the Stat3 protein. The invention relates
CC   to methods for identifying interacting regions of transcription factors
CC   and methods for identifying agents which modulates the interaction
CC   between a transcription factor such as c-Jun and a Stat protein such as
CC   Stat-1 and Stat-3, useful for modulating gene transcription e.g.,
CC   cellular transformation. These identifying agents are used in the
CC   treatment of dysproliferative diseases and also for treating cancer and
CC   psoriasis. A Stat protein comprises the N-terminal domain, coiled-coil
CC   domain, DNA binding domain, linker domain, SH2 domain and transactivation
CC   domain
XX
SQ   Sequence 229 AA;

  Query Match             99.2%;  Score 1160;  DB 4;  Length 229;
  Best Local Similarity   99.1%;  Pred. No. 1.5e-98;
  Matches  227;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||| || ||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHAQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229


RESULT 2
AAY72850
ID   AAY72850 standard; protein; 229 AA.
XX
AC   AAY72850;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 protein fragment #8 (130-358 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .25
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
FT   Region          213. .229
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;  99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 76-77; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 130-358
CC   amino acids of Stat3 protein. This Stat3 fragment showed strong binding
CC   to c-Jun protein in the cell extract. The invention relates to methods
CC   for identifying interacting regions of transcription factors and methods
CC   for identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 229 AA;

  Query Match             98.8%;  Score 1155;  DB 4;  Length 229;
  Best Local Similarity   98.7%;  Pred. No. 4.4e-98;
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||| | | ||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 3
AAY72846
ID   AAY72846 standard; protein; 252 AA.
XX
AC   AAY72846;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 protein fragment #4 (107-358 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          24. .48
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
FT   Region          236. .252
```

```
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 73; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 107-358
CC   amino acids of Stat3 protein. This Stat3 fragment showed strong binding
CC   to c-Jun protein in the cell extract. The invention relates to methods
CC   for identifying interacting regions of transcription factors and methods
CC   for identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 252 AA;

  Query Match             98.8%;  Score 1155;  DB 4;  Length 252;
  Best Local Similarity   98.7%;  Pred. No. 5e-98;
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 83

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 203

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||| | | ||||||||
Db        204 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 252


RESULT 4
AAY72841
ID   AAY72841 standard; protein; 271 AA.
XX
AC   AAY72841;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 protein fragment #2 (107-377 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          24. .48
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
FT   Region          236. .252
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 67-68; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 107-377
CC   amino acids of Stat3 protein. This Stat3 fragment showed strong binding
CC   to c-Jun protein in the cell extract. The invention relates to methods
CC   for identifying interacting regions of transcription factors and methods
CC   for identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 271 AA;

  Query Match             98.8%;  Score 1155;  DB 4;  Length 271;
  Best Local Similarity   98.7%;  Pred. No. 5.5e-98;
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 83

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 203

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||| | | ||||||||
Db        204 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 252


RESULT 5
AAR72082
ID   AAR72082 standard; protein; 770 AA.
XX
AC   AAR72082;
XX
DT   25-MAR-2003  (revised)
DT   27-SEP-1995  (first entry)
XX
DE   Mouse Stat3 (19sf6).
XX
```

```
KW   Signal transducer and activator of transcription; STAT; 19sf6; Stat3;
KW   receptor recognition factor; transcription factor; cellular debilitation;
KW   derangement; dysfunction; interferon-gamma.
XX
OS   Mus sp.
XX
PN   WO9508629-A1.
XX
PD   30-MAR-1995.
XX
PF   26-SEP-1994;   94WO-US010849.
XX
PR   24-SEP-1993;   93US-00126588.
PR   24-SEP-1993;   93US-00126595.
PR   11-MAR-1994;   94US-00212184.
PR   11-MAR-1994;   94US-00212185.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Darnell JE,  Schindler CW,  Shuai K,  Wen Z,  Zhong Z;
XX
DR   WPI; 1995-139598/18.
DR   N-PSDB; AAQ89340.
XX
PT   Receptor recognition factor implicated in transcriptional stimulation of
PT   genes - useful in drug screening assays and/or for treating cellular
PT   debilitations, derangements and/or dysfunctions, etc.
XX
PS   Claim 1; Page 107-110; 160pp; English.
XX
CC   A fragment encoding the human Stat91 protein was used to screen a murine
CC   thymus and spleen cDNA for homologous proteins. A highly homologous gene
CC   (given in AAQ89338) was isolated that encoded a 91 kDa protein (AAR72080)
CC   (Stat1) that was responsive to interferon- gamma. Using a fragment of the
CC   mouse gene as probe, 2 additional members of the 113-91 family of
CC   receptor recognition factor proteins were isolated. The 2 genes (AAQ89339
CC   -40) were cloned in plasmids 13sf1 and 19sf6 and encoded proteins termed
CC   Stat4 (AAR72081) and Stat3 (AAR72082), respectively. (Updated on 25-MAR-
CC   2003 to correct PN field.)
XX
SQ   Sequence 770 AA;

  Query Match             98.8%;  Score 1155;  DB 2;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 2.2e-97;
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 6
AAW03176
ID   AAW03176 standard; protein; 770 AA.
XX
AC   AAW03176;
XX
DT   24-OCT-1996  (first entry)
XX
DE   Mouse STAT4.
XX
KW   STAT; STAT4; signal transducer and activator of transcription;
KW   DNA binding protein; ligand; receptor; oncogenesis; inflammation;
KW   autoimmune disease; antagonist; therapy.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Domain          398. .508
FT                   /label= DNA_binding_domain
FT                   /note= "Claim 3, page 110"
XX
PN   WO9620954-A2.
XX
PD   11-JUL-1996.
XX
PF   28-DEC-1995;   95WO-US017025.
XX
PR   06-JAN-1995;   95US-00369796.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Darnell JE,  Wen Z,  Horvath CM,  Zhong Z;
XX
DR   WPI; 1996-333941/33.
DR   N-PSDB; AAT31280.
XX
PT   New STAT protein DNA-binding domain peptide(s) - useful for diagnosing,
PT   preventing or treating cellular dysfunction, e.g. oncogenesis,
PT   inflammation, parasitic disease or autoimmunity.
XX
PS   Disclosure; Page 87-90; 138pp; English.
XX
CC   Mouse signal transducer and activator of transcription (STAT) protein
CC   STAT4 (AAW03176) serves a dual purpose, i.e. signal transduction from
CC   ligand-activated receptor kinase complexes followed by nuclear
CC   translocation and DNA binding to activate transcription. Recombinant
CC   STAT4 can be obtd. using cDNA clone 19sf6 (AAT31278) obtd. from
CC   splenic/thymic cells. STAT4 includes a DNA-binding domain (see also
CC   AAW03167) capable of both receptor recognition and message delivery via
CC   DNA binding in a receptor-ligand specific manner. STAT proteins and their
CC   DNA binding domains (see also AAW03165-75) are useful for screening
CC   antagonists used to inhibit STAT-mediated signal transduction and
CC   activation of transcription
XX
SQ   Sequence 770 AA;

  Query Match             98.8%;  Score 1155;  DB 2;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 2.2e-97;
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 7
AAY72863
ID   AAY72863 standard; protein; 229 AA.
XX
AC   AAY72863;
XX
```

```
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 mutant (V151A) protein fragment (130-358 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy; mutant; mutein.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .25
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
FT   Misc-difference 22
FT                   /note= "Wild type Val substituted with Ala corresponds to
FT                   151 position of Stat-3 protein"
FT   Region          213. .229
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 66; Page 86; 86pp; English.
XX
CC   The present sequence is mouse Stat3 mutant (V151A) protein fragment
CC   containing 130-358 amino acids of Stat3 protein. This mutant is obtained
CC   by replacing Val 151 with Ala in Stat3 protein. The invention relates to
CC   methods for identifying interacting regions of transcription factors and
CC   methods for identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 229 AA;

  Query Match             98.5%;  Score 1151;  DB 4;  Length 229;
  Best Local Similarity   98.3%;  Pred. No. 1e-97;
  Matches  225;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||| | | ||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
```

```
RESULT 8
AAY72862
ID   AAY72862 standard; protein; 229 AA.
XX
AC   AAY72862;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 mutant (L148A) protein fragment (130-358 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy; mutant; mutein.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .25
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
FT   Misc-difference 19
FT                   /note= "Wild type Leu substituted with Ala; corresponds
FT                   to 148 position of Stat-3 protein"
FT   Region          213. .229
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 66; Page 85; 86pp; English.
XX
CC   The present sequence is mouse Stat3 mutant (L148A) protein fragment
CC   containing 130-358 amino acids of Stat3 protein. This mutant is obtained
CC   by replacing Leu 148 with Ala in Stat3 protein. The invention relates to
CC   methods for identifying interacting regions of transcription factors and
CC   methods for identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 229 AA;

  Query Match             98.4%;  Score 1150;  DB 4;  Length 229;
  Best Local Similarity   98.3%;  Pred. No. 1.3e-97;
  Matches  225;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
```

```
Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||| | | ||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
```

RESULT 9
AAE22055
ID AAE22055 standard; protein; 720 AA.
XX
AC AAE22055;
XX
DT 25-JUL-2002 (first entry)
XX
DE Human Stat3beta protein.
XX
KW Human; signal transducer and activator of transcription 3; ischaemia;
KW immune response; Stat3; coronary atherosclerosis; vascular occlusion;
KW hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;
KW inflammation; chronic obstructive pulmonary disease; cardiac arrest;
KW insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;
KW shock; chronic active hepatitis; adult respiratory distress syndrome;
KW nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;
KW Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;
KW polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;
KW proliferative disease; Grave's disease; ulcerative colitis; sarcoma;
KW carcinoma; degenerative disorder; gene therapy; growth deficiency;
KW cirrhosis; hypoproliferative disorder; lesion; Statbeta.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 713. .714
FT /note= "Encoded by ACA CCA TTC"
XX
PN WO200220032-A1.
XX
PD 14-MAR-2002.
XX
PF 10-SEP-2001; 2001WO-US028254.
XX
PR 08-SEP-2000; 2000US-0231212P.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
PA (UYSF-) UNIV SOUTH FLORIDA.
XX
PI Yu H, Pardoll D, Jove R, Dalton W;
XX
DR WPI; 2002-362218/39.
DR N-PSDB; AAD35066.
XX
PT Modulating angiogenesis and an immune response in an individual, for
PT treating a hypoxic or ischemic condition, comprises administering a
PT compound that modulates the activity of a signal transducer and activator
PT of transcription 3.
XX
PS Disclosure; Page 87-89; 94pp; English.
XX
CC The invention relates to a method of modulating angiogenesis and immune
CC response. Method involves administering to an individual a compound that
CC modulate the activity of signal transducer and activator of transcription
CC 3 (Stat3). Modulating angiogenesis is useful for treating or preventing
CC hypoxic or ischaemic condition or disorder which is the result of stroke,
CC ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,
CC tissue ischaemia in the lower extremities, infarction, trauma, vascular
CC occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,
CC chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,
CC epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest,
CC nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy
CC with neovascularisation. Suppressing an immune response is useful for
CC ameliorating a symptom of an autoimmune disease such as systemic lupus
CC erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,
CC Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,
CC mixed connective tissue disease, primary biliary cirrhosis, pernicious
CC anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,
CC gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,
CC idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's
CC disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune
CC infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and
CC dense deposit disease. The method is useful in preventing or treating
CC specific proliferative and oncogenic disease which includes sarcomas and
CC carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,
CC fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,
CC hypoproliferative disorders, physical trauma, lesions and wounds. The
CC method is also used in gene therapy. The present sequence is human
CC Stat3beta protein
XX
SQ Sequence 720 AA;

Query Match 97.9%; Score 1144; DB 5; Length 720;
Best Local Similarity 97.8%; Pred. No. 2e-96;
Matches 224; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358
```

RESULT 10
ABB57164
ID ABB57164 standard; protein; 769 AA.
XX
AC ABB57164;
XX
DT 07-MAR-2002 (first entry)
XX
DE Mouse ischaemic condition related protein sequence SEQ ID NO:398.
XX
KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease.
XX
OS Mus musculus.
XX
PN WO200188188-A2.
XX
PD 22-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-JP004192.
XX
PR 18-MAY-2000; 2000JP-00145977.
XX
PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
DR WPI; 2002-034733/04.
DR N-PSDB; ABI99454.
XX
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or by

PT determining the expression profile of a gene group comprising these
PT genes.
XX
PS Claim 2; Page 1084-1087; 2690pp; English.
XX
CC The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (ABI99202 to ABI99912, encoding the
CC protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The expression
CC levels or expression profiles produced by these genes are used as an
CC indicator when screening for ischaemic condition-improving drugs or
CC therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR
CC primers for a mouse ischaemic condition related sequence, which are used
CC in the exemplification of the present invention
XX
SQ Sequence 769 AA;

Query Match 97.9%; Score 1144; DB 5; Length 769;
Best Local Similarity 97.8%; Pred. No. 2.2e-96;
Matches 224; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              ||||||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358
```

RESULT 11
AAE22054
ID AAE22054 standard; protein; 769 AA.
XX
AC AAE22054;
XX
DT 25-JUL-2002 (first entry)
XX
DE Human Stat3 protein.
XX
KW Human; signal transducer and activator of transcription 3; ischaemia;
KW immune response; Stat3; coronary atherosclerosis; vascular occlusion;
KW hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;
KW inflammation; chronic obstructive pulmonary disease; cardiac arrest;
KW insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;
KW shock; chronic active hepatitis; adult respiratory distress syndrome;
KW nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;
KW Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;
KW polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;
KW proliferative disease; Grave's disease; ulcerative colitis; sarcoma;
KW carcinoma; degenerative disorder; gene therapy; growth deficiency;
KW cirrhosis; hypoproliferative disorder; lesion.
XX
OS Homo sapiens.
XX
PN WO200220032-A1.
XX
PD 14-MAR-2002.
XX
PF 10-SEP-2001; 2001WO-US028254.
XX
PR 08-SEP-2000; 2000US-0231212P.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
PA (UYSF-) UNIV SOUTH FLORIDA.
XX
PI Yu H, Pardoll D, Jove R, Dalton W;
XX
DR WPI; 2002-362218/39.
DR N-PSDB; AAD35065.
XX
PT Modulating angiogenesis and an immune response in an individual, for
PT treating a hypoxic or ischemic condition, comprises administering a
PT compound that modulates the activity of a signal transducer and activator
PT of transcription 3.
XX
PS Disclosure; Page 83-85; 94pp; English.
XX
CC The invention relates to a method of modulating angiogenesis and immune
CC response. Method involves administering to an individual a compound that
CC modulate the activity of signal transducer and activator of transcription
CC 3 (Stat3). Modulating angiogenesis is useful for treating or preventing
CC hypoxic or ischaemic condition or disorder which is the result of stroke,
CC ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,
CC tissue ischaemia in the lower extremities, infarction, trauma, vascular
CC occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,
CC chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,
CC epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest,
CC nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy
CC with neovascularisation. Suppressing an immune response is useful for
CC ameliorating a symptom of an autoimmune disease such as systemic lupus
CC erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,
CC Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,
CC mixed connective tissue disease, primary biliary cirrhosis, pernicious
CC anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,
CC gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,
CC idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's
CC disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune
CC infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and
CC dense deposit disease. The method is useful in preventing or treating
CC specific proliferative and oncogenic disease which includes sarcomas and
CC carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,
CC fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,
CC hypoproliferative disorders, physical trauma, lesions and wounds. The
CC method is also used in gene therapy. The present sequence is human Stat3
CC protein
XX
SQ Sequence 769 AA;

Query Match 97.9%; Score 1144; DB 5; Length 769;
Best Local Similarity 97.8%; Pred. No. 2.2e-96;
Matches 224; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              ||||||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358
```

RESULT 12
AAE22056
ID AAE22056 standard; protein; 769 AA.

XX
AC AAE22056;
XX
DT 25-JUL-2002 (first entry)
XX
DE Human protein related to angiogenesis regulation.
XX
KW Human; signal transducer and activator of transcription 3; ischaemia;
KW immune response; Stat3; coronary atherosclerosis; vascular occlusion;
KW hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;
KW inflammation; chronic obstructive pulmonary disease; cardiac arrest;
KW insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;
KW shock; chronic active hepatitis; adult respiratory distress syndrome;
KW nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;
KW Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;
KW polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;
KW proliferative disease; Grave's disease; ulcerative colitis; sarcoma;
KW carcinoma; degenerative disorder; gene therapy; growth deficiency;
KW cirrhosis; hypoproliferative disorder; lesion.
XX
OS Homo sapiens.
XX
PN WO200220032-A1.
XX
PD 14-MAR-2002.
XX
PF 10-SEP-2001; 2001WO-US028254.
XX
PR 08-SEP-2000; 2000US-0231212P.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
PA (UYSF-) UNIV SOUTH FLORIDA.
XX
PI Yu H, Pardoll D, Jove R, Dalton W;
XX
DR WPI; 2002-362218/39.
XX
PT Modulating angiogenesis and an immune response in an individual, for
PT treating a hypoxic or ischemic condition, comprises administering a
PT compound that modulates the activity of a signal transducer and activator
PT of transcription 3.
XX
PS Disclosure; Page 83-85; 94pp; English.
XX
CC The invention relates to a method of modulating angiogenesis and immune
CC response. Method involves administering to an individual a compound that
CC modulate the activity of signal transducer and activator of transcription
CC 3 (Stat3). Modulating angiogenesis is useful for treating or preventing
CC hypoxic or ischaemic condition or disorder which is the result of stroke,
CC ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,
CC tissue ischaemia in the lower extremities, infarction, trauma, vascular
CC occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,
CC chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,
CC epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest,
CC nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy
CC with neovascularisation. Suppressing an immune response is useful for
CC ameliorating a symptom of an autoimmune disease such as systemic lupus
CC erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,
CC Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,
CC mixed connective tissue disease, primary biliary cirrhosis, pernicious
CC anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,
CC gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,
CC idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's
CC disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune
CC infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and
CC dense deposit disease. The method is useful in preventing or treating
CC specific proliferative and oncogenic disease which includes sarcomas and
CC carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,
CC fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,
CC hypoproliferative disorders, physical trauma, lesions and wounds. The
CC method is also used in gene therapy. The present sequence is human
CC protein related to angiogenesis regulation
XX

SQ Sequence 769 AA;

Query Match 97.9%; Score 1144; DB 5; Length 769;
Best Local Similarity 97.8%; Pred. No. 2.2e-96;
Matches 224; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358
```

RESULT 13
AAR82995
ID AAR82995 standard; protein; 770 AA.
XX
AC AAR82995;
XX
DT 25-MAR-1996 (first entry)
XX
DE Mouse liver acute phase response factor.
XX
KW Mouse; acute phase response factor; transcription factor; interleukin-6;
KW signal transmission; liver; antibody; antisense; ribozyme;
KW antiinflammatory; antitumor; hypotensive; therapy.
XX
OS Mus musculus.
XX
PN EP676469-A2.
XX
PD 11-OCT-1995.
XX
PF 29-MAR-1995; 95EP-00104670.
XX
PR 04-APR-1994; 94JP-00065825.
XX
PA (KISH/) KISHIMOTO T.
XX
PI Akira S, Kishimoto T;
XX
DR WPI; 1995-346089/45.
DR N-PSDB; AAT05619.
XX
PT New acute phase response factor - for developing inhibitory agents for
PT treating diseases induced by cytokine(s) such as IL-6, e.g. inflammatory
PT diseases.
XX
PS Claim 10; Page 20-22; 31pp; English.
XX
CC The sequence represents a mouse acute phase response factor (APRF), a
CC transcription factor related to signal transmission of interleukin-6 (IL-
CC 6). The protein is encoded by a cDNA, isolated from a mouse liver cDNA
CC library using a polymerase chain reaction product (amplified using
CC primers derived from an IL-6-treated mouse liver peptide) as a probe.
CC APRF-inhibitors, e.g. antibodies, antisense oligonucleotides or
CC ribozymes, may be used to treat diseases induced by IL-6, e.g.
CC inflammatory disease, leukemia, cancer, osteoclasia, pulmonary
CC hypertension, etc
XX
SQ Sequence 770 AA;

Query Match 97.9%; Score 1144; DB 2; Length 770;

```
  Best Local Similarity   97.8%;  Pred. No. 2.2e-96;
  Matches  224;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 14
AAY03768
ID   AAY03768 standard; protein; 770 AA.
XX
AC   AAY03768;
XX
DT   11-JUN-1999  (first entry)
XX
DE   Human STAT3 allelic variant.
XX
KW   Signal Transducer and Activator of Transcription 3; STAT3; allele; IL-6;
KW   intracellular transcription factor; interleukin-6; medicament; variant;
KW   pharmaceutical; autoimmune disease; inflammatory; human.
XX
OS   Homo sapiens.
XX
PN   EP905234-A2.
XX
PD   31-MAR-1999.
XX
PF   18-FEB-1998;   98EP-00102774.
XX
PR   16-SEP-1997;   97EP-00116061.
XX
PA   (ISTF ) ARS APPLIED RES SYSTEMS HOLDING NV.
XX
PI   Serlupi-Crescenzi O,  Della Pietra L;
XX
DR   WPI; 1999-192664/17.
DR   N-PSDB; AAX29281.
XX
PT   New human Signal Transducer and Activator of Transcription 3 (STAT3)
PT   allelic variant useful for treatment of autoimmune and inflammatory
PT   disease.
XX
PS   Claim 2; Page 9-13; 32pp; English.
XX
CC   The present sequence represents a predominant allelic variant of human
CC   Signal Transducer and Activator of Transcription 3 (STAT3) protein, an
CC   intracellular transcription factor which mediates IL-6 signals. The
CC   encoding sequence differs from the original published human STAT3 gene
CC   sequence (Akira et al., 1994). Host cells comprising a recombinant STAT3
CC   DNA molecule can be used for the recombinant expression of the variant.
CC   STAT3 protein is useful as a medicament or pharmaceutical composition for
CC   treatment of autoimmune or inflammatory diseases
XX
SQ   Sequence 770 AA;

  Query Match             97.9%;  Score 1144;  DB 2;  Length 770;
  Best Local Similarity   97.8%;  Pred. No. 2.2e-96;
  Matches  224;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 15
AAB12377
ID   AAB12377 standard; peptide; 770 AA.
XX
AC   AAB12377;
XX
DT   08-NOV-2000  (first entry)
XX
DE   N-terminal domain of murine STAT-3 protein.
XX
KW   STAT; signal transducer and activator of transcription; crystal;
KW   drug design; murine.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Region          4. .9
FT                   /label= Alpha helix 1
FT   Region          12. .21
FT                   /label= Alpha helix 2
FT   Region          19. .21
FT                   /label= 3(10) helix of alpha helix 2
FT   Region          28. .33
FT                   /label= Alpha helix 3
FT   Region          35. .40
FT                   /label= Alpha helix 4
FT   Region          43. .47
FT                   /label= Alpha helix 5
FT   Region          50. .73
FT                   /label= Alpha helix 6
FT   Region          77. .96
FT                   /label= Alpha helix 7
FT   Region          99. .119
FT                   /label= Alpha helix 8
XX
PN   US6087478-A.
XX
PD   11-JUL-2000.
XX
PF   23-JAN-1998;   98US-00012710.
XX
PR   23-JAN-1998;   98US-00012710.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Vinkemeier U,  Moarefi I,  Darnell JE,  Kuriyan J;
XX
DR   WPI; 2000-505108/45.
XX
PT   New crystals of an N-terminal fragment of a signal transducer and
PT   activator of transcription that effectively diffracts x-rays, useful for
PT   drug screening and development.
XX
PS   Disclosure; Fig 1; 42pp; English.
XX
CC   The present invention relates to a crystal of an N-terminal fragment of a
CC   signal transducer and activator of transcription (STAT) protein. The
```

CC crystal effectively diffracts X-rays, allowing the determination of the
CC atomic coordinates of the N-terminal domain to a resolution of greater
CC than 5.0 Angstroms. The present sequence is the N-terminal domain of the
CC murine STAT 3 protein. The N-terminal domain enables STAT dimers to
CC interact and bind DNA cooperatively, a mechanism important for gene
CC activation. The crystals are useful in drug screening and development by
CC selecting a potential drug by performing rational drug design with the 3-
CC dimensional structure determined for the crystal
XX
SQ Sequence 770 AA;

Query Match 97.9%; Score 1144; DB 3; Length 770;
Best Local Similarity 97.8%; Pred. No. 2.2e-96;
Matches 224; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358
```

Search completed: May 25, 2005, 17:36:37
Job time : 123.057 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:25:01 ; Search time 28.6534 Seconds
(without alignments)
596.600 Million cell updates/sec

Title: US-10-090-185-29
Perfect score: 1169
Sequence: 1 NHPTAAVVTEKQQMLEQHLQ.....IKTGVQFATAVALLVKFPEL 229

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1169	100.0	229	3	US-09-387-418A-29	Sequence 29, Appl
2	1155	98.8	229	3	US-09-387-418A-18	Sequence 18, Appl
3	1155	98.8	229	3	US-09-387-418A-28	Sequence 28, Appl
4	1155	98.8	252	3	US-09-387-418A-14	Sequence 14, Appl
5	1155	98.8	271	3	US-09-387-418A-9	Sequence 9, Appli
6	1155	98.8	770	1	US-08-369-796-12	Sequence 12, Appl
7	1155	98.8	770	2	US-08-852-091-12	Sequence 12, Appl
8	1155	98.8	770	2	US-08-820-754-12	Sequence 12, Appl
9	1155	98.8	770	3	US-08-956-652-12	Sequence 12, Appl
10	1155	98.8	770	3	US-08-956-869-12	Sequence 12, Appl
11	1155	98.8	770	3	US-08-948-547-12	Sequence 12, Appl
12	1155	98.8	770	3	US-09-364-970-3	Sequence 3, Appli
13	1155	98.8	770	3	US-09-364-970-5	Sequence 5, Appli
14	1155	98.8	770	3	US-08-956-653A-12	Sequence 12, Appl
15	1155	98.8	770	4	US-08-212-185-12	Sequence 12, Appl
16	1155	98.8	770	5	PCT-US95-17025-12	Sequence 12, Appl
17	1151	98.5	229	3	US-09-387-418A-31	Sequence 31, Appl
18	1150	98.4	229	3	US-09-387-418A-30	Sequence 30, Appl
19	1144	97.9	770	1	US-08-416-581B-9	Sequence 9, Appli
20	1144	97.9	770	3	US-09-012-710-8	Sequence 8, Appli
21	1144	97.9	770	3	US-09-556-273-8	Sequence 8, Appli
22	1144	97.9	770	3	US-09-526-542-2	Sequence 2, Appli
23	1144	97.9	770	4	US-10-117-087-2	Sequence 2, Appli
24	1139	97.4	770	1	US-08-416-581B-1	Sequence 1, Appli
25	1139	97.4	770	1	US-08-416-581B-5	Sequence 5, Appli
26	1139	97.4	770	3	US-09-087-465-6	Sequence 6, Appli
27	1139	97.4	770	4	US-09-972-800A-6	Sequence 6, Appli
28	1139	97.4	771	1	US-08-276-099A-14	Sequence 14, Appl
29	1139	97.4	771	1	US-08-781-890-14	Sequence 14, Appl
30	1094	93.6	213	3	US-09-387-418A-19	Sequence 19, Appl
31	1094	93.6	236	3	US-09-387-418A-15	Sequence 15, Appl
32	1027	87.9	223	3	US-09-387-418A-22	Sequence 22, Appl
33	828	70.8	185	3	US-09-387-418A-23	Sequence 23, Appl
34	781	66.8	176	3	US-09-387-418A-16	Sequence 16, Appl
35	653	55.9	128	3	US-09-387-418A-20	Sequence 20, Appl
36	605	51.8	143	3	US-09-387-418A-17	Sequence 17, Appl
37	556	47.6	268	3	US-09-387-418A-12	Sequence 12, Appl
38	556	47.6	582	4	US-09-430-806A-3	Sequence 3, Appli
39	556	47.6	712	1	US-08-369-796-6	Sequence 6, Appli
40	556	47.6	712	2	US-08-852-091-6	Sequence 6, Appli
41	556	47.6	712	2	US-08-820-754-6	Sequence 6, Appli
42	556	47.6	712	3	US-08-956-652-6	Sequence 6, Appli
43	556	47.6	712	3	US-08-956-869-6	Sequence 6, Appli
44	556	47.6	712	3	US-08-948-547-6	Sequence 6, Appli
45	556	47.6	712	3	US-08-956-653A-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-09-387-418A-29
; Sequence 29, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/09/387,418A
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-387-418A-29

Query Match 100.0%; Score 1169; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 4e-99;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
```

RESULT 2
US-09-387-418A-18
; Sequence 18, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H

```
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 18
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-18

  Query Match             98.8%;  Score 1155;  DB 3;  Length 229;
  Best Local Similarity   98.7%;  Pred. No. 7.6e-98;
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              ||||||||||||||||||||||||||||||||||| | | ||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 3
US-09-387-418A-28
; Sequence 28, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
;  APPLICANT:  Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 28
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-28

  Query Match             98.8%;  Score 1155;  DB 3;  Length 229;
  Best Local Similarity   98.7%;  Pred. No. 7.6e-98;
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              ||||||||||||||||||||||||||||||||||| | | ||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 4
US-09-387-418A-14
; Sequence 14, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
;  APPLICANT:  Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 252
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-14

  Query Match             98.8%;  Score 1155;  DB 3;  Length 252;
  Best Local Similarity   98.7%;  Pred. No. 8.6e-98;
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 83

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 203

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              ||||||||||||||||||||||||||||||||||| | | ||||||||
Db        204 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 252


RESULT 5
US-09-387-418A-9
; Sequence 9, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
;  APPLICANT:  Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 271
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-9

  Query Match             98.8%;  Score 1155;  DB 3;  Length 271;
  Best Local Similarity   98.7%;  Pred. No. 9.4e-98;
```

```
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 83

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 203

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              ||||||||||||||||||||||||||||||||||| | | ||||||||
Db        204 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 252


RESULT 6
US-08-369-796-12
; Sequence 12, Application US/08369796
; Patent No. 5716622
;  GENERAL INFORMATION:
;    APPLICANT:  James E. Darnell, Jr.
;    APPLICANT:  Zilong Wen
;    APPLICANT:  Curt M. Horvath
;    APPLICANT:  Zhong Zhong
;    TITLE OF INVENTION:  FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
;    TITLE OF INVENTION:  TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/369,796
;      FILING DATE:  06-JAN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-116
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-369-796-12

  Query Match             98.8%;  Score 1155;  DB 1;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 3.6e-97;
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              ||||||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 7
US-08-852-091-12
; Sequence 12, Application US/08852091
; Patent No. 5883228
;  GENERAL INFORMATION:
;    APPLICANT:  James E. Darnell, Jr.
;    APPLICANT:  Zilong Wen
;    APPLICANT:  Curt M. Horvath
;    APPLICANT:  Zhong Zhong
;    TITLE OF INVENTION:  FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
;    TITLE OF INVENTION:  TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/852,091
;      FILING DATE:  06-MAY-1997
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/369,796
;      FILING DATE:  06-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-116
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-852-091-12

  Query Match             98.8%;  Score 1155;  DB 2;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 3.6e-97;
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309
```

```
Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              ||||||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 8
US-08-820-754-12
; Sequence 12, Application US/08820754
; Patent No. 5976835
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/820,754
;      FILING DATE:  19-MAR-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,185
;      FILING DATE:  11-MAR-1994
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-820-754-12

  Query Match             98.8%;  Score 1155;  DB 2;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 3.6e-97;
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              ||||||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 9
US-08-956-652-12
; Sequence 12, Application US/08956652
; Patent No. 6013475
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/956,652
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,185
;      FILING DATE:  11-MAR-1994
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX;  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-956-652-12

  Query Match             98.8%;  Score 1155;  DB 3;  Length 770;
```

```
  Best Local Similarity   98.7%;  Pred. No. 3.6e-97;
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 10
US-08-956-869-12
; Sequence 12, Application US/08956869
; Patent No. 6030808
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/956,869
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/212,185
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-956-869-12

  Query Match             98.8%;  Score 1155;  DB 3;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 3.6e-97;
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 11
US-08-948-547-12
; Sequence 12, Application US/08948547
; Patent No. 6124118
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/948,547
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/212,185
;      FILING DATE:  11-MAR-1994
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-948-547-12

  Query Match             98.8%;  Score 1155;  DB 3;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 3.6e-97;
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              ||||||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 12
US-09-364-970-3
; Sequence 3, Application US/09364970
; Patent No. 6235873
; GENERAL INFORMATION:
;  APPLICANT: Bromberg, Jacqueline
;  TITLE OF INVENTION: CONSTITUTIVELY ACTIVE TRANSCRIPTION FACTORS AND THEIR
;  TITLE OF INVENTION: USES FOR IDENTIFYING MODULATORS OF ACTIVITY INCLUDING
;  TITLE OF INVENTION: DYSPROLIFERATIVE CELLULAR CHANGES
;  FILE REFERENCE: 600-1-252
;  CURRENT APPLICATION NUMBER: US/09/364,970
;  CURRENT FILING DATE: 1999-07-31
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-364-970-3

  Query Match             98.8%;  Score 1155;  DB 3;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 3.6e-97;
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              ||||||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 13
US-09-364-970-5
; Sequence 5, Application US/09364970
; Patent No. 6235873
; GENERAL INFORMATION:
;  APPLICANT: Bromberg, Jacqueline
;  TITLE OF INVENTION: CONSTITUTIVELY ACTIVE TRANSCRIPTION FACTORS AND THEIR
;  TITLE OF INVENTION: USES FOR IDENTIFYING MODULATORS OF ACTIVITY INCLUDING
;  TITLE OF INVENTION: DYSPROLIFERATIVE CELLULAR CHANGES
;  FILE REFERENCE: 600-1-252
;  CURRENT APPLICATION NUMBER: US/09/364,970
;  CURRENT FILING DATE: 1999-07-31
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-364-970-5

  Query Match             98.8%;  Score 1155;  DB 3;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 3.6e-97;
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              ||||||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 14
US-08-956-653A-12
; Sequence 12, Application US/08956653A
; Patent No. 6338949
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/956,653A
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,185
;      FILING DATE:  11-MAR-1994
```

```
;      APPLICATION NUMBER: US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-195
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-956-653A-12

  Query Match             98.8%;  Score 1155;  DB 3;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 3.6e-97;
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              ||||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 15
US-08-212-185-12
; Sequence 12, Application US/08212185
; Patent No. 6605442
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/212,185
;      FILING DATE:  11-MAR-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-212-185-12

  Query Match             98.8%;  Score 1155;  DB 4;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 3.6e-97;
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              ||||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


Search completed: May 25, 2005, 17:47:39
Job time : 28.6534 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:36:58 ; Search time 100.287 Seconds
(without alignments)
763.830 Million cell updates/sec

Title: US-10-090-185-29
Perfect score: 1169
Sequence: 1 NHPTAAVVTEKQQMLEQHLQ.....IKTGVQFATAVALLVKFPEL 229

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1434725 seqs, 334507595 residues

Total number of hits satisfying chosen parameters: 1434725

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1169	100.0	229	13	US-10-090-185-29	Sequence 29, Appl
2	1155	98.8	229	13	US-10-090-185-18	Sequence 18, Appl
3	1155	98.8	229	13	US-10-090-185-28	Sequence 28, Appl
4	1155	98.8	252	13	US-10-090-185-14	Sequence 14, Appl
5	1155	98.8	271	13	US-10-090-185-9	Sequence 9, Appli
6	1155	98.8	770	11	US-09-876-773-12	Sequence 12, Appl
7	1155	98.8	770	17	US-10-639-617-12	Sequence 12, Appl
8	1151	98.5	229	13	US-10-090-185-31	Sequence 31, Appl
9	1150	98.4	229	13	US-10-090-185-30	Sequence 30, Appl
10	1144	97.9	720	15	US-10-380-020-4	Sequence 4, Appli
11	1144	97.9	769	15	US-10-380-020-2	Sequence 2, Appli
12	1144	97.9	769	15	US-10-380-020-5	Sequence 5, Appli
13	1144	97.9	770	14	US-10-045-792-8	Sequence 8, Appli
14	1144	97.9	770	14	US-10-038-010-56	Sequence 56, Appl
15	1144	97.9	770	14	US-10-117-087-2	Sequence 2, Appli
16	1144	97.9	793	9	US-09-925-302-780	Sequence 780, App
17	1144	97.9	793	10	US-09-925-302-780	Sequence 780, App
18	1141	97.6	770	15	US-10-116-275-329	Sequence 329, App
19	1139	97.4	770	15	US-10-116-275-349	Sequence 349, App
20	1094	93.6	213	13	US-10-090-185-19	Sequence 19, Appl
21	1094	93.6	236	13	US-10-090-185-15	Sequence 15, Appl
22	1027	87.9	223	13	US-10-090-185-22	Sequence 22, Appl
23	828	70.8	185	13	US-10-090-185-23	Sequence 23, Appl
24	781	66.8	176	13	US-10-090-185-16	Sequence 16, Appl
25	653	55.9	128	13	US-10-090-185-20	Sequence 20, Appl
26	605	51.8	143	13	US-10-090-185-17	Sequence 17, Appl
27	560	47.9	749	9	US-09-833-205-4	Sequence 4, Appli
28	556	47.6	268	13	US-10-090-185-12	Sequence 12, Appl
29	556	47.6	582	14	US-10-245-120-3	Sequence 3, Appli
30	556	47.6	712	11	US-09-876-773-6	Sequence 6, Appli
31	556	47.6	712	14	US-10-245-120-2	Sequence 2, Appli
32	556	47.6	712	17	US-10-639-617-6	Sequence 6, Appli
33	556	47.6	712	17	US-10-936-390-5	Sequence 5, Appli
34	556	47.6	750	9	US-09-833-205-2	Sequence 2, Appli
35	556	47.6	750	11	US-09-876-773-4	Sequence 4, Appli
36	556	47.6	750	14	US-10-245-120-1	Sequence 1, Appli
37	556	47.6	750	14	US-10-308-279-44	Sequence 44, Appl
38	556	47.6	750	16	US-10-755-889-352	Sequence 352, App
39	556	47.6	750	16	US-10-755-889-823	Sequence 823, App
40	556	47.6	750	17	US-10-492-043-19	Sequence 19, Appl
41	556	47.6	750	17	US-10-639-617-4	Sequence 4, Appli
42	556	47.6	786	9	US-09-925-297-550	Sequence 550, App
43	554	47.4	129	13	US-10-090-185-24	Sequence 24, Appl
44	549	47.0	749	9	US-09-833-205-6	Sequence 6, Appli
45	549	47.0	749	11	US-09-876-773-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-10-090-185-29
; Sequence 29, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/10/090,185
; CURRENT FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: 09/387,418
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-090-185-29

Query Match 100.0%; Score 1169; DB 13; Length 229;
Best Local Similarity 100.0%; Pred. No. 1.9e-94;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
```

```
Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229


RESULT 2
US-10-090-185-18
; Sequence 18, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-18

  Query Match             98.8%;  Score 1155;  DB 13;  Length 229;
  Best Local Similarity   98.7%;  Pred. No. 3.2e-93;
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||| | | ||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 3
US-10-090-185-28
; Sequence 28, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-28

  Query Match             98.8%;  Score 1155;  DB 13;  Length 229;
  Best Local Similarity   98.7%;  Pred. No. 3.2e-93;
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||| | | ||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 4
US-10-090-185-14
; Sequence 14, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 252
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-14

  Query Match             98.8%;  Score 1155;  DB 13;  Length 252;
  Best Local Similarity   98.7%;  Pred. No. 3.6e-93;
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 83

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 203

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||| | | ||||||||
Db        204 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 252


RESULT 5
US-10-090-185-9
```

```
; Sequence 9, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 271
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-9

  Query Match             98.8%;  Score 1155;  DB 13;  Length 271;
  Best Local Similarity   98.7%;  Pred. No. 3.9e-93;
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 83

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 203

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              ||||||||||||||||||||||||||||||||||| | | ||||||||
Db        204 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 252


RESULT 6
US-09-876-773-12
; Sequence 12, Application US/09876773
; Publication No. US20040058318A1
;  GENERAL INFORMATION:
;       APPLICANT: Darnell Jr., James E.
;                  Schindler, Christian W.
;                  Fu, Xian-Yuan
;                  Wen, Zilong
;                  Zhong, Zhong
;       TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
;                           SEQUENCES AND METHODS OF USE THEREOF
;       NUMBER OF SEQUENCES: 25
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Klauber & Jackson
;            STREET: 411 Hackensack Avenue
;            CITY: Hackensack
;            STATE: New Jersey
;            COUNTRY: USA
;            ZIP: 07601
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/876,773
;            FILING DATE: 07-Jun-2001
;            CLASSIFICATION: 435
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/212,185
;            FILING DATE: 11-MAR-1994
;            APPLICATION NUMBER: US 07/980,498
;            FILING DATE: 23-NOV-1992
;            APPLICATION NUMBER: US 07/854,296
;            FILING DATE: 19-MAR-1992
;            APPLICATION NUMBER: WO US93/02569
;            FILING DATE: 19-MAR-1993
;            APPLICATION NUMBER: US 08/126,588
;            FILING DATE: 24-SEP-1993
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Jackson Esq., David A.
;            REGISTRATION NUMBER: 26,742
;            REFERENCE/DOCKET NUMBER: 600-1-073 CIP
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 201 487-5800
;            TELEFAX: 201 343-1684
;            TELEX: 133521
;  INFORMATION FOR SEQ ID NO: 12:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 770 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-876-773-12

  Query Match             98.8%;  Score 1155;  DB 11;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 1.5e-92;
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              ||||||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 7
US-10-639-617-12
; Sequence 12, Application US/10639617
; Publication No. US20050079543A1
;  GENERAL INFORMATION:
;       APPLICANT: Darnell Jr., James E.
;                  Schindler, Christian W.
;                  Fu, Xian-Yuan
;                  Wen, Zilong
;                  Zhong, Zhong
;       TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
;                           SEQUENCES AND METHODS OF USE THEREOF
;       NUMBER OF SEQUENCES: 25
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Klauber & Jackson
;            STREET: 411 Hackensack Avenue
;            CITY: Hackensack
;            STATE: New Jersey
;            COUNTRY: USA
;            ZIP: 07601
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/639,617
;            FILING DATE: 12-Aug-2003
;            CLASSIFICATION: 435
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/212,185
;            FILING DATE: 11-MAR-1994
;            APPLICATION NUMBER: US 07/980,498
;            FILING DATE: 23-NOV-1992
;            APPLICATION NUMBER: US 07/854,296
;            FILING DATE: 19-MAR-1992
;            APPLICATION NUMBER: WO US93/02569
;            FILING DATE: 19-MAR-1993
;            APPLICATION NUMBER: US 08/126,588
;            FILING DATE: 24-SEP-1993
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Jackson Esq., David A.
;            REGISTRATION NUMBER: 26,742
;            REFERENCE/DOCKET NUMBER: 600-1-073 CIP
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 201 487-5800
;            TELEFAX: 201 343-1684
;            TELEX: 133521
;   INFORMATION FOR SEQ ID NO: 12:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 770 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-639-617-12

  Query Match             98.8%;  Score 1155;  DB 17;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 1.5e-92;
  Matches  226;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 8
US-10-090-185-31
; Sequence 31, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-31

  Query Match             98.5%;  Score 1151;  DB 13;  Length 229;
  Best Local Similarity   98.3%;  Pred. No. 7.1e-93;
  Matches  225;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||| | | ||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 9
US-10-090-185-30
; Sequence 30, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-30

  Query Match             98.4%;  Score 1150;  DB 13;  Length 229;
  Best Local Similarity   98.3%;  Pred. No. 8.7e-93;
  Matches  225;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||| | | ||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 10
US-10-380-020-4
; Sequence 4, Application US/10380020
```

```
; Publication No. US20040052762A1
; GENERAL INFORMATION:
;  APPLICANT: Yu, Hua
;  APPLICANT:  Pardoll, Drew
;  APPLICANT:  Jove, Richard
;  APPLICANT:  Dalton, Willian
;  TITLE OF INVENTION: Stat3 Agonists and Antagonists and Therapeutic Uses Thereof
;  FILE REFERENCE: 10873-009-999
;  CURRENT APPLICATION NUMBER: US/10/380,020
;  CURRENT FILING DATE:  2003-03-07
;  PRIOR APPLICATION NUMBER: 60/231,212
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
;   LENGTH: 720
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-380-020-4

  Query Match            97.9%;  Score 1144;  DB 15;  Length 720;
  Best Local Similarity   97.8%;  Pred. No. 1.3e-91;
  Matches  224;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 11
US-10-380-020-2
; Sequence 2, Application US/10380020
; Publication No. US20040052762A1
; GENERAL INFORMATION:
;  APPLICANT: Yu, Hua
;  APPLICANT:  Pardoll, Drew
;  APPLICANT:  Jove, Richard
;  APPLICANT:  Dalton, Willian
;  TITLE OF INVENTION: Stat3 Agonists and Antagonists and Therapeutic Uses Thereof
;  FILE REFERENCE: 10873-009-999
;  CURRENT APPLICATION NUMBER: US/10/380,020
;  CURRENT FILING DATE:  2003-03-07
;  PRIOR APPLICATION NUMBER: 60/231,212
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 769
;   TYPE: PRT<
;   ORGANISM: Homo Sapiens
US-10-380-020-2

  Query Match            97.9%;  Score 1144;  DB 15;  Length 769;
  Best Local Similarity   97.8%;  Pred. No. 1.4e-91;
  Matches  224;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 12
US-10-380-020-5
; Sequence 5, Application US/10380020
; Publication No. US20040052762A1
; GENERAL INFORMATION:
;  APPLICANT: Yu, Hua
;  APPLICANT:  Pardoll, Drew
;  APPLICANT:  Jove, Richard
;  APPLICANT:  Dalton, Willian
;  TITLE OF INVENTION: Stat3 Agonists and Antagonists and Therapeutic Uses Thereof
;  FILE REFERENCE: 10873-009-999
;  CURRENT APPLICATION NUMBER: US/10/380,020
;  CURRENT FILING DATE:  2003-03-07
;  PRIOR APPLICATION NUMBER: 60/231,212
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 769
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-380-020-5

  Query Match            97.9%;  Score 1144;  DB 15;  Length 769;
  Best Local Similarity   97.8%;  Pred. No. 1.4e-91;
  Matches  224;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 13
US-10-045-792-8
; Sequence 8, Application US/10045792
; Publication No. US20030003563A1
;   GENERAL INFORMATION:
;        APPLICANT: Vinkemeier, Uwe
;                   Moarefi, Ismail
;                   Darnell, Jr., James E.
;                   Kuriyan, John
;        TITLE OF INVENTION: A CRYSTAL OF THE N-TERMINAL DOMAIN OF A
;                            STAT PROTEIN AND METHODS OF USE THEREOF
;        NUMBER OF SEQUENCES: 13
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Klauber & Jackson
;             STREET: 411 Hackensack Avenue, 4th Floor
;             CITY: Hackensack
;             STATE: New Jersey
```

```
;             COUNTRY: USA
;             ZIP: 07601
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/045,792
;             FILING DATE: 19-Oct-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/012,710
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Jackson Esq., David A.
;             REGISTRATION NUMBER: 26,742
;             REFERENCE/DOCKET NUMBER: 600-1-194
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 201-487-5800
;             TELEFAX: 201-343-1684
;             TELEX: 133521
;   INFORMATION FOR SEQ ID NO: 8:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 770 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        HYPOTHETICAL: NO
;        SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-045-792-8

  Query Match             97.9%;  Score 1144;  DB 14;  Length 770;
  Best Local Similarity   97.8%;  Pred. No. 1.4e-91;
  Matches  224;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 14
US-10-038-010-56
; Sequence 56, Application US/10038010
; Publication No. US20030040089A1
; GENERAL INFORMATION:
;  APPLICANT: HYBRIGENICS
;  APPLICANT:  Pierre, Legrain
;  TITLE OF INVENTION: Protein-protein interactions in adipocyte cells
;  FILE REFERENCE: B4767A
;  CURRENT APPLICATION NUMBER: US/10/038,010
;  CURRENT FILING DATE:  2002-07-23
;  PRIOR APPLICATION NUMBER: US 60/259,377
;  PRIOR FILING DATE: 2001-01-02
;  NUMBER OF SEQ ID NOS: 67
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: STAT3 : Transcription factor
;   LOCATION: (1)..(770)
;   OTHER INFORMATION:
US-10-038-010-56

  Query Match             97.9%;  Score 1144;  DB 14;  Length 770;
  Best Local Similarity   97.8%;  Pred. No. 1.4e-91;
  Matches  224;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 15
US-10-117-087-2
; Sequence 2, Application US/10117087
; Publication No. US20030166854A1
; GENERAL INFORMATION:
;  APPLICANT: SERLUPI-CRESCENZI, Ottaviano
;  APPLICANT:  DELLA PIETRA, Linda
;  TITLE OF INVENTION: ALLELIC VARIANTS OF HUMAN STAT3
;  FILE REFERENCE: SERLUPI=2
;  CURRENT APPLICATION NUMBER: US/10/117,087
;  CURRENT FILING DATE:  2002-04-08
;  PRIOR APPLICATION NUMBER: US/09/526,542
;  PRIOR FILING DATE: 2000-03-19
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Human
US-10-117-087-2

  Query Match             97.9%;  Score 1144;  DB 14;  Length 770;
  Best Local Similarity   97.8%;  Pred. No. 1.4e-91;
  Matches  224;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


Search completed: May 25, 2005, 18:21:53
Job time : 101.62 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:24:07 ; Search time 21.6038 Seconds
(without alignments)
1019.898 Million cell updates/sec

Title: US-10-090-185-29
Perfect score: 1169
Sequence: 1 NHPTAAVVTEKQQMLEQHLQ.....IKTGVQFATAVALLVKFPEL 229

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1144	97.9	770	2	I49508	ISGF3 p91-related
2	1139	97.4	770	2	A54444	DNA-binding protei
3	510.5	43.7	739	2	A46159	interferon-depende
4	495	42.3	748	2	A56047	gamma-interferon a
5	310	26.5	851	2	A46160	interferon alpha-i
6	269.5	23.1	786	2	I49274	mammary gland fact
7	269.5	23.1	793	2	S54772	mammary gland fact
8	263.5	22.5	794	2	G02317	transcription acti
9	238	20.4	794	2	S55527	mammary gland fact
10	126	10.8	978	2	A70387	conserved hypothet
11	124	10.6	533	2	G72593	hypothetical prote
12	122	10.4	848	2	A54740	interleukin-4-indu
13	120.5	10.3	837	2	I57557	DNA-Binding Protei
14	119.5	10.2	1166	2	T27075	hypothetical prote
15	118.5	10.1	1208	2	AE1947	chromosome segrega
16	118.5	10.1	2094	2	S33124	tpr protein - huma
17	115.5	9.9	924	2	S06117	myosin heavy chain
18	115	9.8	1164	2	T24806	hypothetical prote
19	114.5	9.8	2007	1	B43402	myosin heavy chain
20	114	9.8	1976	2	A59252	myosin heavy chain
21	113.5	9.7	727	2	AC1814	hypothetical prote
22	113.5	9.7	857	2	S33821	median body protei
23	112.5	9.6	946	2	S28061	SCP1 protein - rat
24	112.5	9.6	1690	2	T13030	microtubule bindin
25	112.5	9.6	1957	2	T38077	hypothetical coile
26	112.5	9.6	2253	2	T30336	nuclear/mitotic ap
27	112	9.6	289	2	S51193	epimorphin - mouse
28	112	9.6	1509	1	A27224	myosin heavy chain
29	112	9.6	1999	1	S21801	myosin heavy chain
30	111.5	9.5	284	2	C64527	M protein - Helico
31	111.5	9.5	1048	1	BVECSC	exonuclease (EC 3.
32	111	9.5	1959	1	A33977	myosin heavy chain
33	110.5	9.5	747	1	A57107	kinesin-related pr
34	110.5	9.5	1940	2	A59287	myosin heavy chain
35	110.5	9.5	1992	2	A47297	myosin heavy chain
36	110	9.4	1188	2	G83960	chromosome segrega
37	110	9.4	2020	2	T21174	hypothetical prote
38	109.5	9.4	434	2	T43448	hypothetical prote
39	109.5	9.4	527	2	S33068	myosin heavy chain
40	109.5	9.4	734	2	T27055	hypothetical prote
41	109.5	9.4	896	2	S43074	epidermal growth f
42	109.5	9.4	1047	2	C85535	ATP-dependent dsDN
43	109.5	9.4	1047	2	G90684	ATP-dependent dsDN
44	109.5	9.4	1738	2	T14867	interaptin - slime
45	109.5	9.4	3685	1	A27605	dystrophin, muscle

ALIGNMENTS

RESULT 1
I49508
ISGF3 p91-related transcription factor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I49508; I49009
R;Akira, S.; Nishio, Y.; Inoue, M.; Wang, X.J.; Wei, S.; Matsusaka, T.; Yoshida, K.; Su
Cell 77, 63-71, 1994
A;Title: Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related tr
A;Reference number: A54444; MUID:94208062; PMID:7512451
A;Accession: I49508
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-770 <RES>
A;Cross-references: UNIPROT:P42227; GB:L29278; NID:g476715; PIDN:AAA37254.1; PID:g47671
R;Raz, R.; Durbin, J.E.; Levy, D.E.
J. Biol. Chem. 269, 24391-24395, 1994
A;Title: Acute phase response factor and additional members of the interferon-stimulate
A;Reference number: I49009; MUID:95014185; PMID:7523373
A;Accession: I49009
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-393,'M',395-700,702-770 <RE2>
A;Cross-references: EMBL:U08378; NID:g473889; PIDN:AAA56668.1; PID:g473890
C;Genetics:
A;Gene: APRF
C;Superfamily: human signal transducer and transcription activator STAT5A

Query Match 97.9%; Score 1144; DB 2; Length 770;
Best Local Similarity 97.8%; Pred. No. 1e-68;
Matches 224; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358
```

RESULT 2
A54444
DNA-binding protein APRF - human

```
C;Species: Homo sapiens (man)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 09-Jul-2004
C;Accession: A54444
R;Akira, S.; Nishio, Y.; Inoue, M.; Wang, X.J.; Wei, S.; Matsusaka, T.; Yoshida, K.; Sud
Cell 77, 63-71, 1994
A;Title: Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related tra
A;Reference number: A54444; MUID:94208062; PMID:7512451
A;Accession: A54444
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-770 <RES>
A;Cross-references: UNIPROT:P40763; GB:L29277; NID:g475788; PID:g475789
C;Genetics:
A;Gene: GDB:STAT3; APRF
A;Cross-references: GDB:358950
A;Map position: 17q21-17q21
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: DNA binding; transcription factor

  Query Match             97.4%;  Score 1139;  DB 2;  Length 770;
  Best Local Similarity   97.4%;  Pred. No. 2.3e-68;
  Matches  223;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELHQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 3
A46159
interferon-dependent positive-acting transcription factor ISGF-3 91K chain - human
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A46159
R;Schindler, C.; Fu, X.Y.; Improta, T.; Aebersold, R.; Darnell Jr., J.E.
Proc. Natl. Acad. Sci. U.S.A. 89, 7836-7839, 1992
A;Title: Proteins of transcription factor ISGF-3: one gene encodes the 91-and 84-kDa ISG
A;Reference number: A46159; MUID:92366557; PMID:1502203
A;Accession: A46159
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid; protein
A;Residues: 1-739 <SCH>
A;Cross-references: UNIPROT:P42224
A;Experimental source: HeLa cells
A;Note: sequence extracted from NCBI backbone (NCBIP:110818)
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             43.7%;  Score 510.5;  DB 2;  Length 739;
  Best Local Similarity   44.9%;  Pred. No. 1.4e-26;
  Matches  101;  Conservative   49;  Mismatches   62;  Indels   13;  Gaps    2;

Qy          5 AAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQ 64
              : |: :||: |:  :::|: :|  :| ::| :|:|||::||  |||  |    : ||  :
Db        132 STVMLDKQKELDSKVRNVKDKVMCIEHEIKSLEDLQDEYDFKCKTL--QNREHETNGVAK 189

Qy         65 SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEIACIG 124
              |  :|:   |::|   ||  |: :| ::  ||:  |  |  | ::|| :|||| : ||||
Db        190 SDQKQEQLLLKKMYLMLDNKRKEVVHKIIELLNVTELTQNALINDELVEWKRRQQSACIG 249

Qy        125 GPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERIVELF 184
              |||| |||:|:           | |||:||||||:|| :|: ||| ::: :| :|   ||
Db        250 GPPNACLDQLQ----------QVRQQLKKLEELEQKYTYEHDPITKNKQVLWDRTFSLF 298

Qy        185 RNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              : |::|:|||||||||| || ||||:||||||   : ||||  ||
Db        299 QQLIQSSFVVERQPCMPTHPQRPLVLKTGVQFTVKLRLLVKLQEL 343


RESULT 4
A56047
gamma-interferon activation site-binding protein Stat4 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004
C;Accession: A56047
R;Yamamoto, K.; Quelle, F.W.; Thierfelder, W.E.; Kreider, B.L.; Gilbert, D.J.; Jenkins,
Mol. Cell. Biol. 14, 4342-4349, 1994
A;Title: Stat4, a novel gamma interferon activation site-binding protein expressed in e
A;Reference number: A56047; MUID:94277038; PMID:8007943
A;Accession: A56047
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-748 <YAM>
A;Cross-references: UNIPROT:P42228; GB:U09351; NID:g509502; PIDN:AAA19692.1; PID:g50950
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: DNA binding; phosphoprotein

  Query Match             42.3%;  Score 495;  DB 2;  Length 748;
  Best Local Similarity   44.2%;  Pred. No. 1.6e-25;
  Matches   99;  Conservative   49;  Mismatches   68;  Indels    8;  Gaps    2;

Qy          8 VTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKS--QGDMQDLNGNNQS 65
              |:|:|: :|  :  ::  ||  ||  | :|:|||:||: |||:::  |||        |
Db        136 VSERQRNVEHKVSAIKNSVQMTEQDTKYLEDLQDEFDYRYKTIQTMDQGD------KNSI 189

Qy         66 VTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEIACIGG 125
              :  |::  |::|| :||  |:  :|::  :::  : :  ::  ||| |||:| :||||||
Db        190 LVNQEVLTLQEMLNSLDFKRKEALSKMTQIVNETDLLMNSMLLEELQDWKKRQQIACIGG 249

Qy        126 PPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERIVELFR 185
              | :  ||:|:|  | ||||  | |||::||:|   |::|:||||   |  | ||   |
Db        250 PLHNGLDQLQNCFTLLAESLFQLRQQLEKLQEQSTKMTYEGDPIPAQRAHLLERATFLIY 309

Qy        186 NLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              || |::|||||||||| || ||:|:|| :||   : ||:| |||
Db        310 NLFKNSFVVERQPCMPTHPQRPMVLKTLIQFTVKLRLLIKLPEL 353


RESULT 5
A46160
interferon alpha-induced transcription activator ISGF-3, 113K chain - human
N;Alternate names: stat2 protein
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A46160; S71908; S53873
R;Fu, X.Y.; Schindler, C.; Improta, T.; Aebersold, R.; Darnell Jr., J.E.
Proc. Natl. Acad. Sci. U.S.A. 89, 7840-7843, 1992
A;Title: The proteins of ISGF-3, the interferon alpha-induced transcriptional activator
A;Reference number: A46160; MUID:92366558; PMID:1502204
A;Accession: A46160
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA; protein
A;Residues: 1-851 <FU1>
A;Cross-references: UNIPROT:P52630
A;Note: sequence extracted from NCBI backbone (NCBIP:110820)
R;Yan, R.; Qureshi, S.; Zhong, Z.; Wen, Z.; Darnell, J.E.
submitted to the EMBL Data Library, December 1994
A;Reference number: S71908
A;Accession: S71908
A;Molecule type: DNA
A;Residues: 1-851 <YAN>
A;Cross-references: EMBL:U18671; NID:g1293919; PIDN:AAA98760.1; PID:g1293920
R;Yan, R.; Qureshi, S.; Zhong, Z.; Wen, Z.; Darnell Jr., J.E.
Nucleic Acids Res. 23, 459-463, 1995
A;Title: The genomic structure of the STAT genes: multiple exons in coincident sites in
```

```
A;Reference number: S53873; MUID:95192056; PMID:7885841
A;Accession: S53873
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-196;392-591;684-730 <YAW>
A;Cross-references: EMBL:U18671
C;Genetics:
A;Gene: stat2
A;Introns: 44/2; 95/3; 127/3; 157/3; 183/1; 211/3; 261/2; 314/2; 345/2; 365/2; 372/2; 40
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: signal transduction; transcription regulation

  Query Match             26.5%;  Score 310;  DB 2;  Length 851;
  Best Local Similarity   32.6%;  Pred. No. 3.7e-13;
  Matches   72;  Conservative   51;  Mismatches   92;  Indels    6;  Gaps    4;

Qy          8 VTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVT 67
              |  :|  :|  : |:|  :: | : :  ::: || | | || ::::|    |: :
Db        138 VESQQHEIESRILDLRAMMEKLVKSISQLKDQQDVFCFRYK-IQAKGKTPSLDPH--QTK 194

Qy         68 RQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEIACIGGPP 127
               ||:  |:: |  ||: |: ::     ||  :  : : |   :| :|| : : |||  |
Db        195 EQKI--LQETLNELDKRRKEVLDASKALLGRLTTLIELLL-PKLEEWKAQQQKACIRAPI 251

Qy        128 NICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNL 187
              :  |::|| | |: |:     || :|:|: |   |||: ||: :   :   :: || : |
Db        252 DHGLEQLETWFTAGAKLLFHLRQLLKELKGLSCLVSYQDDPLTKGVDLRNAQVTELLQRL 311

Qy        188 MKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPE 228
              :  ||||| |||||  | |||::||| :|     |||:  |
Db        312 LHRAFVVETQPCMPQTPHRPLILKTGSKFTVRTRLLVRLQE 352


RESULT 6
I49274
mammary gland factor - mouse
N;Alternate names: STAT5 protein homolog p80
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I49274; S54773; S54727
R;Liu, X.; Robinson, G.W.; Gouilleux, F.; Groner, B.; Hennighausen, L.
Proc. Natl. Acad. Sci. U.S.A. 92, 8831-8835, 1995
A;Title: Cloning and expression of Stat5 and an additional homologue (Stat5b) involved i
A;Reference number: I49273; MUID:96004632; PMID:7568026
A;Accession: I49274
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-786 <RES>
A;Cross-references: UNIPROT:P42232; UNIPROT:Q9JKM1; EMBL:U21110; NID:g747973; PIDN:AAC52
R;Mui, A.L.F.; Wakao, H.; O'Farrell, A.M.; Harada, N.; Miyajima, A.
EMBO J. 14, 1166-1175, 1995
A;Title: Interleukin-3, granulocyte-macrophage colony stimulating factor and interleukin
A;Reference number: S54772; MUID:95237198; PMID:7720707
A;Accession: S54773
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-432,'E',434-786 <MUI>
A;Cross-references: EMBL:Z48539; NID:g758635; PIDN:CAA88420.1; PID:g758636
R;Azam, M.; Erdjument-Bromage, H.; Kreider, B.L.; Xia, M.; Quelle, F.; Basu, R.; Saris,
EMBO J. 14, 1402-1411, 1995
A;Title: Interleukin-3 signals through multiple isoforms of Stat5.
A;Reference number: S54725; MUID:95246733; PMID:7537213
A;Accession: S54727
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-432,'E',434-786 <AZA>
C;Genetics:
A;Gene: Stat5b
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             23.1%;  Score 269.5;  DB 2;  Length 786;
  Best Local Similarity   30.6%;  Pred. No. 1.7e-10;
  Matches   71;  Conservative   42;  Mismatches   98;  Indels   21;  Gaps    5;

Qy          5 AAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNY-KTLKSQGDMQDLNGNN 63
              |  :::|   : |  :::|   || | ::| ::  |: |   | ::|: |      |  |
Db        134 ADAMSQKHLQINQTFEELRLITQDTENELKKLQQTQEYFIIQYQESLRIQAQFAQLGQLN 193

Qy         64 -------QSVTRQKMQQLEQML----TALDQMRRSIVSELAGLLSAMEYVQKTLTDEELA 112
                     ::  :||   ||  |       | | |  :  :    |  :   |  : |:||
Db        194 PQERMSRETALQQKQVSLETWLQREAQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELI 253

Qy        113 DWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQH 172
              |||| ::|  ||||   || |::|   |||   | ||||:: | | |::   | |: :
Db        254 QWKRRQQLAGNGGPPEGSLDVLQSWCEKLAEIIWQNRQQIRRAEHLCQQLPIPG-PVEEM 312

Qy        173 RPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLV 224
                 :   | ::   |: | |::|:||        | |:||  :||  | |||
Db        313 LAEVNATITDIISALVTSTFIIEKQP--------PQVLKTQTKFAATVRLLV 356


RESULT 7
S54772
mammary gland factor - mouse
N;Alternate names: stat5 protein
C;Species: Mus musculus (house mouse)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S54772; I49273
R;Mui, A.L.F.; Wakao, H.; O'Farrell, A.M.; Harada, N.; Miyajima, A.
EMBO J. 14, 1166-1175, 1995
A;Title: Interleukin-3, granulocyte-macrophage colony stimulating factor and interleuki
A;Reference number: S54772; MUID:95237198; PMID:7720707
A;Accession: S54772
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-793 <MUI>
A;Cross-references: UNIPROT:P42230; UNIPROT:Q9JIA0; EMBL:Z48538; NID:g758633; PIDN:CAA8
R;Liu, X.; Robinson, G.W.; Gouilleux, F.; Groner, B.; Hennighausen, L.
Proc. Natl. Acad. Sci. U.S.A. 92, 8831-8835, 1995
A;Title: Cloning and expression of Stat5 and an additional homologue (Stat5b) involved
A;Reference number: I49273; MUID:96004632; PMID:7568026
A;Accession: I49273
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-793 <RES>
A;Cross-references: EMBL:U21103; NID:g747971; PIDN:AAA80590.1; PID:g747972
C;Genetics:
A;Gene: Stat5a
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             23.1%;  Score 269.5;  DB 2;  Length 793;
  Best Local Similarity   30.6%;  Pred. No. 1.7e-10;
  Matches   70;  Conservative   42;  Mismatches   96;  Indels   21;  Gaps    5;

Qy          8 VTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNY-KTLKSQGDMQDLNGNN--- 63
              :::|   : |  :::|   || | ::| ::  |: |   | ::|: |      |  |
Db        137 MSQKHLQINQRFEELRLITQDTENELKKLQQTQEYFIIQYQESLRIQAQFAQLGQLNPQE 196

Qy         64 ----QSVTRQKMQQLEQML----TALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWK 115
                  ::  :||   ||  |       | | |  :  :    |  :   |  : |:||  ||
Db        197 RMSRETALQQKQVSLETWLQREAQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELIQWK 256

Qy        116 RRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPM 175
              || ::|  ||||   || |::|   |||   | ||||:: | | |::   | |: :
Db        257 RRQQLAGNGGPPEGSLDVLQSWCEKLAEIIWQNRQQIRRAEHLCQQLPIPG-PVEEMLAE 315

Qy        176 LEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLV 224
              :   | ::   |: | |::|:||        | |:||  :||  | |||
Db        316 VNATITDIISALVTSTFIIEKQP--------PQVLKTQTKFAATVRLLV 356


RESULT 8
G02317
transcription activator stat5A - human
C;Species: Homo sapiens (man)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 09-Jul-2004
```

```
C;Accession: G02317
R;Lin, J.
submitted to the EMBL Data Library, December 1995
A;Reference number: H01043
A;Accession: G02317
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-794 <LIN>
A;Cross-references: UNIPROT:P42229; EMBL:U43185; NID:g1151169; PIDN:AAB06589.1; PID:g115
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             22.5%;  Score 263.5;  DB 2;  Length 794;
  Best Local Similarity   30.1%;  Pred. No. 4.3e-10;
  Matches   69;  Conservative   43;  Mismatches   96;  Indels   21;  Gaps    5;

Qy          8 VTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNY-KTLKSQGDMQDL------- 59
              :::|   : |  :::|   || | ::| ::  |: |   | ::|: |     |
Db        137 MSQKHLQINQTFEELRLVTQDTENELKKLQQTQEYFIIQYQESLRIQAQFAQLAQLSPQE 196

Qy         60 NGNNQSVTRQKMQQLEQML----TALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWK 115
               : ::  :||   || |       | | |  :  :    |  :   |  : |:||  ||
Db        197 RLSRETALQQKQVSLEAWLQREAQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELIQWK 256

Qy        116 RRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPM 175
              || ::|  ||||   || |::|   |||   | ||||:: | |.|::   | |: :
Db        257 RRQQLAGNGGPPEGSLDVLQSWCEKLAEIIWQNRQQIRRAEHLCQQLPIPG-PVEEMLAE 315

Qy        176 LEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLV 224
              :   | ::   |: | |::|:||       | |:||  :||  | |||
Db        316 VNATITDIISALVTSTFIIEKQP--------PQVLKTQTKFAATVRLLV 356


RESULT 9
S55527
mammary gland factor - sheep
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S55527; S44353
R;Wakao, H.; Gouilleux, F.; Groner, B.
EMBO J. 14, 854-855, 1995
A;Title: Corrigenda. Mammary gland factor (MGF) is a novel member of the cytokine regula
A;Reference number: S55527; MUID:95188889; PMID:7882987
A;Accession: S55527
A;Molecule type: mRNA
A;Residues: 1-794 <WAK>
A;Cross-references: UNIPROT:P42231; EMBL:X78428; NID:g602354; PIDN:CAA55191.1; PID:g6023
A;Note: this is a revision to the sequence from reference S44353
R;Wakao, H.; Gouilleux, F.; Groner, B.
EMBO J. 13, 2182-2191, 1994
A;Title: Mammary gland factor (MGF) is a novel member of the cytokine regulated transcri
A;Reference number: S44353; MUID:94244619; PMID:7514531
A;Accession: S44353
A;Molecule type: mRNA
A;Residues: 1-716,'RHLHGPGSLPSR',729,'P',731,'ASL' <WAW>
A;Cross-references: EMBL:X78428
A;Note: this sequence has been revised in reference S55527
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             20.4%;  Score 238;  DB 2;  Length 794;
  Best Local Similarity   29.3%;  Pred. No. 2.1e-08;
  Matches   67;  Conservative   42;  Mismatches   98;  Indels   22;  Gaps    6;

Qy          8 VTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNY-KTLKSQGDMQDLNGNN--- 63
              :::|   : |  :::|   || | ::| ::  |: |   | ::|: |     |   |
Db        138 MSQKHLQINQTFEELRLVTQDTENELKKLQQTQEYFIIQYQESLRIQAQFAQLAQLNPQE 197

Qy         64 ----QSVTRQKMQQLEQML----TALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWK 115
                 ::  :||   || |       | | |  :  :    |  :   |  : |:||  ||
Db        198 RLSRETALQQKQVSLEAWLQREAQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELIQWK 257

Qy        116 RRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPM 175
              || :   :  ||   || |::|   |||   | ||||:: | | |::   | |: :
Db        258 RRHDWRGMEAPPR-SLDVLQSWCEKLAEIIWQNRQQIRRAEHLCQQLPIPG-PVEEMLAE 315

Qy        176 LEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLV 224
              :   | ::   |: | |::|:||       | |:||  :||  | |||
Db        316 VNATITDIISALVTSTFIIEKQP--------PQVLKTQTKFAATVRLLV 356


RESULT 10
A70387
conserved hypothetical protein aq_1006 - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C;Accession: A70387
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: A70387
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-978 <AQF>
A;Cross-references: UNIPROT:O67124; GB:AE000718; NID:g2983504; PIDN:AAC07092.1; PID:g29
A;Experimental source: strain VF5
C;Genetics:
A;Gene: aq_1006
C;Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

  Query Match             10.8%;  Score 126;  DB 2;  Length 978;
  Best Local Similarity   23.8%;  Pred. No. 0.77;
  Matches   50;  Conservative   38;  Mismatches   80;  Indels   42;  Gaps    6;

Qy         10 EKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQ 69
              |::  ||: |  |  ::::||   | || |::  :|: |                :
Db        234 EEKDSLERELSQVVTKLKELENLEKEVEKLREKLEFSRKVAP----------YVPIAK 281

Qy         70 KMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLT--------------DEELAD- 113
              ::::::: || |   :  :  ||| |   : : |: |                ::||
Db        282 RIEEIDKKLTELKVRKNKLTKELAVLKDELSFAQEELNRIEAEKEKFKEEKEREKELEHR 341

Qy        114 WKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHR 173
               |:  ||  |       |  |    :|| | : :  |  :: |:| ::|  ||  :|
Db        342 LKKLQEIKEI-------LKELSQLSSSLKEKEREYEQAKQEFEDLSERVE-KGKKLVAET 393

Qy        174 PMLEERIVELFRN------LMKSAFVVERQ 197
                 |:| |||         ||   :|| |
Db        394 EEKLEKIKELFSEEEYTSLKMKERLLVELQ 423


RESULT 11
G72593
hypothetical protein APE1216 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: G72593
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Taka
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: G72593
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-533 <KAW>
A;Cross-references: UNIPROT:Q9YCP2; DDBJ:AP000061; NID:g5104821; PIDN:BAA80205.1; PID:d
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1216

  Query Match             10.6%;  Score 124;  DB 2;  Length 533;
  Best Local Similarity   21.2%;  Pred. No. 0.52;
  Matches   42;  Conservative   49;  Mismatches   65;  Indels   42;  Gaps    7;

Qy          8 VTEKQQMLEQHLQDVRKRVQDLEQKMKVVEN----LQDDF-------------DFNYK 48
```

```
Db        319 MSQQLQALAEDLESLSSRVEDLEARVGSVEDRLSQAEEDIDSLTTSLDSLRTELEDLSTR 378

Qy         49 TLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTD 108

Db        379 LAEAQASLEDLNTRLDQVA-STLQQLQQRLATAEESLQALTEDLASLQAEVETLQQSIVE 437

Qy        109 EELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQ----TRQQIKKLEELQQKVSY 164

Db        438 IDRRLGQLRSTVDAV---------RLE--VESLGEKLVQAEEKNQRQDASIEDFQSQIQ- 485

Qy        165 KGDPIVQHRPMLEERIVE 182

Db        486 ------ELRTQLDEKTRE 497


RESULT 12
A54740
interleukin-4-induced transcription factor stat - human
C;Species: Homo sapiens (man)
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004
C;Accession: A54740
R;Hou, J.; Schindler, U.; Henzel, W.J.; Ho, T.C.; Brasseur, M.; McKnight, S.L.
Science 265, 1701-1706, 1994
A;Title: An interleukin-4-induced transcription factor: IL-4 stat.
A;Reference number: A54740; MUID:94367369; PMID:8085155
A;Accession: A54740
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-848 <HOU>
A;Cross-references: UNIPROT:P42226
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: DNA binding; transcription regulation

  Query Match             10.4%;  Score 122;  DB 2;  Length 848;
  Best Local Similarity   24.3%;  Pred. No. 1.2;
  Matches   58;  Conservative   37;  Mismatches   94;  Indels   50;  Gaps    9;

Qy          3 PTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGN 62

Db         92 PLKLVAT-----FRQILQGKKKAV------MEQFRHLPMPFHWKQEELKFKTGLRRLQHR 140

Qy         63 NQSV--TRQKMQQ--------LEQML------TALDQMRRSIVSELAGLLSAMEYVQKTL 106

Db        141 VGEIHLLREALQKGAEAGQVSLHSLIETPANGTGPSEALAMLLQETTGELEA----AKAL 196

Qy        107 TDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKG 166

Db        197 VLKRIQIWKRQQQLAGNGAP----------FEESLAPLQERCESLVDIYSQLQQEVGAAG 246

Qy        167 DPI-VQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLV 224

Db        247 GELEPKTRASLTGRLDEVLRTLVTSCFLVEKQP--------PQVLKTQTKFQAGVRFLL 297


RESULT 13
I57557
DNA-Binding Protein and transcription factor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C;Accession: I57557
R;Quelle, F.W.; Shimoda, K.; Thierfelder, W.; Fischer, C.L.; Kim, A.; Ruben, S.M.; Cleve
Mol. Cell. Biol. 15, 3336-3343, 1995
A;Title: Cloning of murine Stat6 and human Stat6, Stat proteins that are tyrosine phosph
A;Reference number: I57557; MUID:95280934; PMID:7760829
A;Accession: I57557
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-837 <RES>
A;Cross-references: UNIPROT:P52633; GB:L47650; NID:g1008876; PIDN:AAA79006.1; PID:g10088
C;Genetics:
A;Gene: STAT6
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             10.3%;  Score 120.5;  DB 2;  Length 837;
  Best Local Similarity   25.8%;  Pred. No. 1.5;
  Matches   60;  Conservative   31;  Mismatches   83;  Indels   59;  Gaps   10;

Qy         10 EKQQMLEQ--HLQDVRKRVQD----------LEQKMKVVENLQDDFDFNYKTLKSQGDMQ 57

Db        106 EKKAVIEEFRHLPGPFHRKQEELKFTTPLGRLHHRVRETRLLRESLHLGPKT--GQVSLQ 163

Qy         58 D-----LNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELA 112

Db        164 NLIDPPLNGPGPS---------EDLPTILQ----------GTVGDLETTQ-PLVLLRIQ 202

Qy        113 DWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLE-ELQQKVSYKGDPIVQ 171

Db        203 IWKRQQQLAGNGTPFEESLAGLQERCESLVEIYSQLHQEIGAASGELEPKT--------- 253

Qy        172 HRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLV 224

Db        254 -RASLISRLDEVLRTLVTSSFLVEKQP--------PQVLKTQTKFQAGVRFLL 297


RESULT 14
T27075
hypothetical protein Y51A2D.16 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T27075
R;McMurray, A.
submitted to the EMBL Data Library, January 1998
A;Reference number: Z20307
A;Accession: T27075
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1166 <WIL>
A;Cross-references: EMBL:AL021497; PIDN:CAA16403.1; GSPDB:GN00023; CESP:Y51A2D.16
A;Experimental source: clone Y51A2D
C;Genetics:
A;Gene: CESP:Y51A2D.16
A;Map position: 5
A;Introns: 17/3; 45/1; 76/2; 156/3; 240/2; 288/2; 594/2; 657/3; 756/3; 798/2; 900/3; 94

  Query Match             10.2%;  Score 119.5;  DB 2;  Length 1166;
  Best Local Similarity   21.4%;  Pred. No. 2.6;
  Matches   52;  Conservative   46;  Mismatches   72;  Indels   73;  Gaps    9;

Qy          1 NHPTAAVVTEKQ----QMLEQHLQD-------------VRKRVQDLEQKMKVVENLQDD 42

Db        430 SHVTRSLSSEKNTLKAKLLE--LQDQVEAQTLELLNQKNCGKRLEDRDQMISNLHNLKNE 487

Qy         43 FDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEY- 101

Db        488 LENDLKTCQTQLELES----------KKLQRLREDLVLEKSRRADLIGRIHSLCTTLSLN 537

Qy        102 ---VQKTLTDEELAD-----------WKRRPEIACIGGPPNI----------------- 129

Db        538 GANFEKINNDDELIDNIDDIMMNALVAVKRERDDLRIQGNQQIQELHDLKRDIEKLRRSE 597

Qy        130 --CLDRLENWITSLAESQLQTRQQI----KKLEELQQKVSYKGDPIVQHRPMLEERIVEL 183

Db        598 SESLNESDDRVRELTRENMHTKEQVFMLQEKLRELNLELSTKNDEI----DMVKASIEEL 653

Qy        184 FRN 186

Db        654 NRN 656


RESULT 15
AE1947
chromosome segregation protein [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
```

C;Accession: AE1947
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AE1947
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1208 <KUR>
A;Cross-references: UNIPROT:Q8YXT3; GB:BA000019; PIDN:BAB73085.1; PID:g17130474; GSPDB:G
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr1128
C;Superfamily: chromosome segregation protein SMC1

Query Match 10.1%; Score 118.5; DB 2; Length 1208;
Best Local Similarity 23.9%; Pred. No. 3.1;
Matches 55; Conservative 44; Mismatches 82; Indels 49; Gaps 9;

```
Qy          5 AAVVTEKQQM---------LEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLK---S 52
              | : |::||:          || |::: : | |:::::  :    ::      |    :
Db        835 ATIKTQEQQIQQRETAFREAEQRLKNLENQQQRLQERIQEAQQRITEYQTQQTTCTEAIN 894

Qy         53 QGDMQDLNGNNQ-SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEEL 111
                :    |    | | : || |: :||| | |  | | :|  |:   |   : ::  :   |
Db        895 RVSQQTTTINAQITQTRAKLSELEQHLGAEKQKRDTIEQEVRSHLLRQQQLEWEIQKLEE 954

Qy        112 ADWKRRPEIACIGG---------------PPNICLDRLENWITSLAESQLQTRQQI--- 152
                 ||| ::  :                  |  : |: |:  : |||: :||  : :
Db        955 TQLKRREDLTALQSQLQELVPELPNPLPEVPDKVDLEELQKELRSLAK-RLQAMEPVNML 1013

Qy        153 ---------KKLEELQQKV-SYKGDPIVQHRPMLEERIVELFRNLMKSAF 192
                     |:|||| ||: : :|:     |  |  || | |  | : ||
Db       1014 ALEEYERTQKRLEELSQKLQTLEGE-----RTELLLRI-ENFTTLRQIAF 1057
```

Search completed: May 25, 2005, 17:45:28
Job time : 22.6038 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:15:30 ; Search time 95.284 Seconds
(without alignments)
1230.701 Million cell updates/sec

Title: US-10-090-185-29
Perfect score: 1169
Sequence: 1 NHPTAAVVTEKQQMLEQHLQ.....IKTGVQFATAVALLVKFPEL 229

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1144	97.9	770	1	STA3_HUMAN	P40763 homo sapien
2	1144	97.9	770	1	STA3_MOUSE	P42227 mus musculu
3	1141	97.6	770	1	STA3_RAT	P52631 rattus norv
4	1137	97.3	770	1	STA3_BOVIN	P61635 bos taurus
5	1119	95.7	771	2	Q6DV79	Q6dv79 gallus gall
6	1086	92.9	769	2	Q9PVX8	Q9pvx8 xenopus lae
7	1083	92.6	766	2	Q7ZXK3	Q7zxk3 xenopus lae
8	998	85.4	414	2	Q7ZTS5	Q7zts5 brachydanio
9	998	85.4	765	2	Q6DVF3	Q6dvf3 oryzias lat
10	998	85.4	785	2	Q6GUE7	Q6gue7 oryzias lat
11	998	85.4	786	2	Q6NV46	Q6nv46 brachydanio
12	998	85.4	806	2	O93599	O93599 brachydanio
13	974	83.3	764	2	Q90Y16	Q90y16 tetraodon f
14	964	82.5	767	2	O13133	O13133 oncorhynchu
15	639	54.7	163	2	Q9N145	Q9n145 macaca mula
16	562	48.1	751	2	Q8JGN0	Q8jgn0 xenopus lae
17	560	47.9	712	2	Q6P6Q7	Q6p6q7 rattus norv
18	560	47.9	749	2	Q9QXK0	Q9qxk0 rattus norv
19	558	47.7	749	2	Q8C497	Q8c497 mus musculu
20	557	47.6	712	2	Q99K94	Q99k94 mus musculu
21	557	47.6	749	2	Q8C3V4	Q8c3v4 mus musculu
22	557	47.6	749	2	Q9D323	Q9d323 mus musculu
23	557	47.6	755	2	Q8C8M3	Q8c8m3 mus musculu
24	556	47.6	750	1	STA1_HUMAN	P42224 homo sapien
25	556	47.6	750	2	Q68D00	Q68d00 homo sapien
26	556	47.6	757	2	Q764M5	Q764m5 sus scrofa
27	549	47.0	749	1	STA1_MOUSE	P42225 mus musculu
28	532.5	45.6	108	2	Q704W6	Q704w6 bos taurus
29	525.5	45.0	754	2	O13131	O13131 oncorhynchu
30	519	44.4	749	2	O93598	O93598 brachydanio
31	519	44.4	749	2	Q6P943	Q6p943 brachydanio
32	518	44.3	528	2	Q8JFU8	Q8jfu8 brachydanio
33	516.5	44.2	718	2	Q801Y2	Q801y2 carassius a
34	507	43.4	748	1	STA4_HUMAN	Q14765 homo sapien
35	504.5	43.2	758	2	Q90Y17	Q90y17 tetraodon f
36	502.5	43.0	754	2	O13132	O13132 oncorhynchu
37	500	42.8	657	2	Q8AW24	Q8aw24 brachydanio
38	491.5	42.0	749	1	STA4_MOUSE	P42228 mus musculu
39	491	42.0	651	2	Q7ZZ53	Q7zz53 brachydanio
40	491	42.0	667	2	Q8AW20	Q8aw20 brachydanio
41	490.5	42.0	652	2	Q7ZZ77	Q7zz77 brachydanio
42	490	41.9	553	2	Q8JFU7	Q8jfu7 brachydanio
43	490	41.9	748	2	Q66HB2	Q66hb2 rattus norv
44	488	41.7	1153	2	Q8JFS5	Q8jfs5 brachydanio
45	478	40.9	737	2	Q90Y15	Q90y15 tetraodon f

ALIGNMENTS

RESULT 1
STA3_HUMAN
ID STA3_HUMAN STANDARD; PRT; 770 AA.
AC P40763; O14916; Q9BW54;
DT 01-FEB-1995 (Rel. 31, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Signal transducer and activator of transcription 3 (Acute-phase
DE response factor).
GN Name=STAT3; Synonyms=APRF;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Placenta;
RX MEDLINE=94208062; PubMed=7512451; DOI=10.1016/0092-8674(94)90235-6;
RA Akira S., Nishio Y., Inoue M., Wang X.-J., Wei S., Matsusaka T.,
RA Yoshida K., Sudo T., Naruto M., Kishimoto T.;
RT "Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-
RT related transcription factor involved in the gp130-mediated signaling
RT pathway.";
RL Cell 77:63-71(1994).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=98296260; PubMed=9630560; DOI=10.1016/S0378-1119(98)00185-1;
RA Della Pietra L., Bressan A., Pezzotti A., Serlupi-Crescenzi O.;
RT "Highly conserved amino-acid sequence between murine STAT3 and a
RT revised human STAT3 sequence.";
RL Gene 213:119-124(1998).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM DEL-701), AND VARIANT ILE-143.
RA Rieder M.J., Daniels R.L., da Ponte S.H., Hastings N.C., Ahearn M.O.,
RA Rajkumar N., Yi Q., Nickerson D.A.;
RT "SeattleSNPs. NHLBI HL66682 program for genomic applications, UW-
RT FHCRC, Seattle, WA (URL: http://pga.gs.washington.edu).";
RL Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND DEL-701).
RC TISSUE=Kidney, and Pancreas;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,

```
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   SEQUENCE OF 564-704 FROM N.A.
RC   TISSUE=Liver;
RA   Della Pietra L., Bressan A., Pezzotti A.R., Serlupi-Crescenzi O.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   PHOSPHORYLATION ON SERINE.
RX   MEDLINE=95215843; PubMed=7701321;
RA   Zhang X., Blenis J., Li H.-C., Schindler C., Chen-Kiang S.;
RT   "Requirement of serine phosphorylation for formation of STAT-promoter
RT   complexes.";
RL   Science 267:1990-1994(1995).
RN   [7]
RP   INTERACTION WITH NCOA1.
RX   PubMed=11773079; DOI=10.1074/jbc.M111486200;
RA   Giraud S., Bienvenu F., Avril S., Gascan H., Heery D.M., Coqueret O.;
RT   "Functional interaction of STAT3 transcription factor with the
RT   coactivator NcoA/SRC1a.";
RL   J. Biol. Chem. 277:8004-8011(2002).
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes.
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P40763-1; Sequence=Displayed;
CC       Name=Del-701;
CC         IsoId=P40763-2; Sequence=VSP_010474;
CC   -!- TISSUE SPECIFICITY: Heart, brain, placenta, lung, liver, skeletal
CC       muscle, kidney and pancreas.
CC   -!- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,
CC       LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation
CC       is important for the formation of stable DNA-binding STAT3
CC       homodimers and maximal transcriptional activity.
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L29277; AAA58374.1; -.
DR   EMBL; AJ012463; CAA10032.1; -.
DR   EMBL; AY572796; AAS66986.1; -.
DR   EMBL; BC000627; AAH00627.1; -.
DR   EMBL; BC014482; AAH14482.1; -.
DR   EMBL; AF029311; AAB84254.1; -.
DR   PIR; A54444; A54444.
DR   HSSP; P42227; 1BG1.
DR   TRANSFAC; T01493; -.
DR   Genew; HGNC:11364; STAT3.
DR   H-InvDB; HIX0013840; -.
DR   MIM; 102582; -.
DR   GO; GO:0005737; C:cytoplasm; TAS.
DR   GO; GO:0005634; C:nucleus; TAS.
DR   GO; GO:0005062; F:hematopoietin/interferon-class (D200-domain. . .; TAS.
DR   GO; GO:0003700; F:transcription factor activity; TAS.
DR   GO; GO:0006928; P:cell motility; TAS.
DR   GO; GO:0007259; P:JAK-STAT cascade; TAS.
DR   GO; GO:0000122; P:negative regulation of transcription from P. . .; TAS.
DR   GO; GO:0007399; P:neurogenesis; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
KW   Activator; Alternative splicing; DNA-binding; Nuclear protein;
KW   Phosphorylation; Polymorphism; SH2 domain; Transcription regulation.
FT   DOMAIN      580    670       SH2.
FT   MOD_RES     705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES     727    727       Phosphoserine (By similarity).
FT   VARSPLIC    701    701       Missing (in isoform Del-701).
FT                                /FTId=VSP_010474.
FT   VARIANT      32     32       Q -> K (in dbSNP:1803125).
FT                                /FTId=VAR_018683.
FT   VARIANT     143    143       M -> I.
FT                                /FTId=VAR_018679.
FT   CONFLICT    288    288       Q -> H (in Ref. 1).
FT   CONFLICT    460    460       P -> S (in Ref. 1).
FT   CONFLICT    548    548       K -> N (in Ref. 1).
FT   CONFLICT    561    561       F -> Y (in Ref. 1).
FT   CONFLICT    667    667       V -> L (in Ref. 1).
FT   CONFLICT    730    730       T -> A (in Ref. 1).
SQ   SEQUENCE   770 AA;  88067 MW;  6C00632211C8012D CRC64;

  Query Match             97.9%;  Score 1144;  DB 1;  Length 770;
  Best Local Similarity   97.8%;  Pred. No. 4.3e-62;
  Matches  224;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 2
STA3_MOUSE
ID   STA3_MOUSE     STANDARD;      PRT;   770 AA.
AC   P42227;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Signal transducer and activator of transcription 3 (Acute-phase
DE   response factor).
GN   Name=Stat3; Synonyms=Aprf;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A), AND SEQUENCE OF 154-158; 181-185
RP   AND 632-640.
RC   STRAIN=BALB/c; TISSUE=Liver;
RX   MEDLINE=94208062; PubMed=7512451; DOI=10.1016/0092-8674(94)90235-6;
```

```
RA   Akira S., Nishio Y., Inoue M., Wang X.J., Shi W., Matsusaka T.,
RA   Yoshida K., Sudo T., Naruto M., Kishimoto T.;
RT   "Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-
RT   related transcription factor involved in the gp130-mediated signaling
RT   pathway.";
RL   Cell 77:63-71(1994).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   TISSUE=Thymus;
RX   MEDLINE=94188718; PubMed=8140422;
RA   Zhong Z., Wen Z., Darnell J.E. Jr.;
RT   "Stat3: a STAT family member activated by tyrosine phosphorylation in
RT   response to epidermal growth factor and interleukin-6.";
RL   Science 264:95-98(1994).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   TISSUE=Brain;
RX   MEDLINE=95014185; PubMed=7523373;
RA   Raz R., Durbin J.E., Levy D.E.;
RT   "Acute phase response factor and additional members of the interferon-
RT   stimulated gene factor 3 family integrate diverse signals from
RT   cytokines, interferons, and growth factors.";
RL   J. Biol. Chem. 269:24391-24395(1994).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3B).
RC   STRAIN=BALB/c, and C57BL/6; TISSUE=Liver;
RX   MEDLINE=96016116; PubMed=7568080;
RA   Schaefer T.S., Sanders L.K., Nathans D.;
RT   "Cooperative transcriptional activity of Jun and Stat3 beta, a short
RT   form of Stat3.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:9097-9101(1995).
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   STRAIN=129/SvJ;
RX   PubMed=11161808; DOI=10.1006/geno.2000.6433;
RA   Miyoshi K., Cui Y., Riedlinger G., Lehoczky J., Zon L., Oka T.,
RA   Dewar K., Hennighausen L.;
RT   "Structure of the mouse stat 3/5 locus: evolution from Drosophila to
RT   zebrafish to mouse.";
RL   Genomics 71:150-155(2001).
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   STRAIN=C57BL/6J, and NOD/LtJ;
RA   Davoodi-Semiromi A., She J.-X.;
RT   "A mutant Stat5b with weaker DNA binding defines a key defective
RT   pathway in non-obese diabetic (NOD) mice.";
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   STRAIN=FVB/N; TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [8]
RP   PHOSPHORYLATION SITE SER-727, AND MUTAGENESIS.
RX   MEDLINE=95354205; PubMed=7543024; DOI=10.1016/0092-8674(95)90311-9;
RA   Wen Z., Zhong Z., Darnell J.E. Jr.;
RT   "Maximal activation of transcription by Stat1 and Stat3 requires both
RT   tyrosine and serine phosphorylation.";
RL   Cell 82:241-250(1995).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 136-716.
RX   MEDLINE=98334373; PubMed=9671298; DOI=10.1038/28101;
RA   Becker S., Groner B., Mueller C.W.;
RT   "Three-dimensional structure of the Stat3beta homodimer bound to
RT   DNA.";
RL   Nature 394:145-151(1998).
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes. STAT3B interacts with the N-terminal
CC       part of JUN to activate such promoters in a cooperative way.
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=Stat3A;
CC         IsoId=P42227-1; Sequence=Displayed;
CC       Name=Stat3B;
CC         IsoId=P42227-2; Sequence=VSP_006287;
CC       Name=Del-701;
CC         IsoId=P42227-3; Sequence=VSP_010475;
CC   -!- TISSUE SPECIFICITY: STAT3A is seen in the liver, spleen, and
CC       kidney. STAT3B is also detected in the liver, although in a much
CC       less abundant manner.
CC   -!- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,
CC       LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation
CC       is important for the formation of stable DNA-binding STAT3
CC       homodimers and maximal transcriptional activity (By similarity).
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L29278; AAA37254.1; -.
DR   EMBL; U06922; AAA19452.1; -.
DR   EMBL; U08378; AAA56668.1; -.
DR   EMBL; U30709; AAC52612.1; -.
DR   EMBL; AF246978; AAL59017.1; -.
DR   EMBL; AY299489; AAQ75418.1; -.
DR   EMBL; AY299490; AAQ75419.1; -.
DR   EMBL; BC003806; AAH03806.1; -.
DR   PIR; I49508; I49508.
DR   PDB; 1BG1; X-ray; A=1-722.
DR   TRANSFAC; T01574; -.
DR   MGD; MGI:103038; Stat3.
DR   GO; GO:0005737; C:cytoplasm; IDA.
DR   GO; GO:0005634; C:nucleus; IDA.
DR   GO; GO:0005886; C:plasma membrane; IDA.
DR   GO; GO:0003677; F:DNA binding; IDA.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   GO; GO:0016563; F:transcriptional activator activity; IDA.
DR   GO; GO:0007259; P:JAK-STAT cascade; IDA.
DR   GO; GO:0006357; P:regulation of transcription from Pol II pro. . .; IDA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
```

```
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
KW   3D-structure; Activator; Acute phase; Alternative splicing;
KW   Direct protein sequencing; DNA-binding; Nuclear protein;
KW   Phosphorylation; SH2 domain; Transcription regulation.
FT   DOMAIN      580    670       SH2.
FT   MOD_RES     705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES     727    727       Phosphoserine.
FT   VARSPLIC    716    770       TTCSNTIDLPMSPRTLDSLMQFGNNGEGAEPSAGGQFESLT
FT                                FDMDLTSECATSPM -> FIDAVWK (in isoform
FT                                Stat3B).
FT                                /FTId=VSP_006287.
FT   VARSPLIC    701    701       Missing (in isoform Del-701).
FT                                /FTId=VSP_010475.
FT   MUTAGEN     727    727       S->A: Decreased transcriptional
FT                                activation.
FT   CONFLICT     16     16       E -> K (in Ref. 2).
FT   CONFLICT     25     25       S -> T (in Ref. 2 and 4).
FT   CONFLICT    394    394       M -> I (in Ref. 1).
FT   HELIX       139    180
FT   TURN        181    182
FT   TURN        197    198
FT   HELIX       199    237
FT   TURN        238    238
FT   HELIX       239    251
FT   TURN        252    253
FT   HELIX       261    290
FT   TURN        294    295
FT   TURN        297    301
FT   HELIX       302    320
FT   STRAND      321    328
FT   TURN        330    331
FT   TURN        333    334
FT   TURN        336    337
FT   STRAND      338    340
FT   TURN        341    342
FT   STRAND      345    351

  Query Match             97.9%;  Score 1144;  DB 1;  Length 770;
  Best Local Similarity   97.8%;  Pred. No. 4.3e-62;
  Matches  224;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 3
STA3_RAT
ID   STA3_RAT       STANDARD;      PRT;   770 AA.
AC   P52631;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Signal transducer and activator of transcription 3.
GN   Name=Stat3;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=96102059; PubMed=8530402; DOI=10.1074/jbc.270.50.29998;
RA   Ripperger J.A., Fritz S., Richter K., Hocke G.M., Lottspeich F.,
RA   Fey G.H.;
RT   "Transcription factors Stat3 and Stat5b are present in rat liver
RT   nuclei late in an acute phase response and bind interleukin-6 response
RT   elements.";
RL   J. Biol. Chem. 270:29998-30006(1995).
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes.
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation (By similarity).
CC   -!- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,
CC       LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation
CC       is important for the formation of stable DNA-binding STAT3
CC       homodimers and maximal transcriptional activity (By similarity).
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X91810; CAA62920.1; -.
DR   HSSP; P42227; 1BG1.
DR   RGD; 3772; Stat3.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
KW   Activator; DNA-binding; Nuclear protein; Phosphorylation; SH2 domain;
KW   Transcription regulation.
FT   DOMAIN      580    670       SH2.
FT   MOD_RES     705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES     727    727       Phosphoserine (By similarity).
SQ   SEQUENCE   770 AA;  88039 MW;  D74A0C76954754ED CRC64;

  Query Match             97.6%;  Score 1141;  DB 1;  Length 770;
  Best Local Similarity   97.4%;  Pred. No. 6.5e-62;
  Matches  223;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |:|||||||||||||||||||||||||||||| | | ||||||||
Db        310 VDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 4
```

```
STA3_BOVIN
ID   STA3_BOVIN     STANDARD;      PRT;   770 AA.
AC   P61635;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Signal transducer and activator of transcription 3.
GN   Name=STAT3;
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Mammary gland;
RA   Seyfert H.M., Wheeler T.T., Moolenaar A., Pitra C.;
RT   "The STAT5B-encoding gene was flipped across the STAT3/STAT5A-locus
RT   during ruminant evolution.";
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes (By similarity).
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation (By similarity).
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ620655; CAF06182.1; -.
DR   PROSITE; PS50001; SH2; 1.
KW   Activator; DNA-binding; Nuclear protein; Phosphorylation; SH2 domain;
KW   Transcription regulation.
FT   DOMAIN      580    670       SH2.
FT   MOD_RES     705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES     727    727       Phosphoserine (By similarity).
SQ   SEQUENCE   770 AA;  87974 MW;  9CEB147C73E83274 CRC64;

  Query Match             97.3%;  Score 1137;  DB 1;  Length 770;
  Best Local Similarity   97.4%;  Pred. No. 1.1e-61;
  Matches  223;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLLDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||| | | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 5
Q6DV79
ID   Q6DV79      PRELIMINARY;      PRT;   771 AA.
AC   Q6DV79;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Signal transducer and activator of transcription 3.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zhou G.Y., Leung F.C.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY641397; AAT64887.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   771 AA;  88174 MW;  71AC855C5DEC03E2 CRC64;

  Query Match             95.7%;  Score 1119;  DB 2;  Length 771;
  Best Local Similarity   96.1%;  Pred. No. 1.5e-60;
  Matches  219;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          2 HPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNG 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        131 HPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNG 190

Qy         62 NNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEIA 121
              |||||||||||||||||||||||| |||||||||||||| | |||||||||| :||
Db        191 NNQSVTRQKMQQLEQMLTALDQMRRGIVSELAGLLSAMEYVQKMLADEELADWKRRQQIA 250

Qy        122 CIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERIV 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 CIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERIV 310

Qy        182 ELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||| |||||| | | ||||||||
Db        311 ELFRNLMKSAFVVERQPCMPMHPDRPLVFKTGVQFTTKVRLLVKFPEL 358


RESULT 6
Q9PVX8
ID   Q9PVX8      PRELIMINARY;      PRT;   769 AA.
AC   Q9PVX8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat 3.
GN   Name=stat 3;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20107399; PubMed=10642787; DOI=10.1006/dbio.1999.9518;
RA   Nishinakamura R., Matsumoto Y., Matsuda T., Ariizumi T., Heike T.,
RA   Asashima M., Yokota T.;
RT   "Activation of Stat3 by cytokine receptor gp130 ventralizes Xenopus
RT   embryos independent of BMP-4.";
RL   Dev. Biol. 216:481-490(1999).
DR   EMBL; AB017701; BAA86061.1; -.
```

```
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   769 AA;  87974 MW;  0905C03263303069 CRC64;

  Query Match             92.9%;  Score 1086;  DB 2;  Length 769;
  Best Local Similarity   91.7%;  Pred. No. 1.5e-58;
  Matches  210;  Conservative    9;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              :|| |||||||||||||||||||:|||||||||||||||||||||||||||| |: :||
Db        130 SHPNAAVVTEKQQMLEQHLQDVRKKVQDLEQKMKVVENLQDDFDFNYKTLKSQSDLSELN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||:||:|:|:|:|| |||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQLRRTIISDLASLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQIRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||    | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTNKVRLLVKFPEL 358


RESULT 7
Q7ZXK3
ID   Q7ZXK3      PRELIMINARY;      PRT;   766 AA.
AC   Q7ZXK3;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat3-A protein.
GN   Name=stat3-A;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC044717; AAH44717.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   766 AA;  87599 MW;  31018A3321CCEB9C CRC64;

  Query Match             92.6%;  Score 1083;  DB 2;  Length 766;
  Best Local Similarity   90.8%;  Pred. No. 2.3e-58;
  Matches  208;  Conservative   11;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              :|| |||:|||||||||||||:|||||||||||||||||||||||||| |: :||
Db        130 SHPNAAVITEKQQMLEQHLQDVRKKVQDLEQKMKVVENLQDDFDFNYKTLKSQSDLSELN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||:||:|:|:|| |||||||||| |||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQLRRTIISDLASLLSAMEYVQKNLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              |||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIRKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              |||||||||||||||||||||||||||||||    | ||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTNKVRLLVKFPEL 358


RESULT 8
Q7ZTS5
ID   Q7ZTS5      PRELIMINARY;      PRT;   414 AA.
AC   Q7ZTS5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat3 protein.
GN   Name=stat3;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB; TISSUE=Whole body;
```

```
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB; TISSUE=Whole body;
RA   Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC045276; AAH45276.1; -.
DR   HSSP; P42227; 1BG1.
DR   ZFIN; ZDB-GENE-980526-68; stat3.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
SQ   SEQUENCE   414 AA;  48253 MW;  0FFD1B509B7526BD CRC64;

  Query Match             85.4%;  Score 998;  DB 2;  Length 414;
  Best Local Similarity   83.5%;  Pred. No. 1.9e-53;
  Matches  192;  Conservative   21;  Mismatches   15;  Indels    2;  Gaps    2;

Qy          2 HPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLN 60
              |||  |||||||:|| :|||:||||||:|||||::|||||||||||||||||:: ||||
Db        130 HPTGTVVTEKQQILEHNLQDIRKRVQDMEQKMKMLENLQDDFDFNYKTLKSAGELSQDLN 189

Qy         61 GNNQ-SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPE 119
              ||:| : ||||| ||||||:||||:|| ||:|:|||||||::||| |||||||||||| :
Db        190 GNSQAAATRQKMSQLEQMLSALDQLRRQIVTEMAGLLSAMDFVQKNLTDEELADWKRRQQ 249

Qy        120 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 179
              |||||||||||||||| |||||||||| |||||:|||||||||||||:|||| |||:
Db        250 IACIGGPPNICLDRLETWITSLAESQLQIRQQIRKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy        180 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              ||:|||||||||||||||||||||||||||||| | | ||||||||
Db        310 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 359


RESULT 9
Q6DVF3
ID   Q6DVF3      PRELIMINARY;      PRT;   765 AA.
AC   Q6DVF3;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Signal transducer and activation of transcription factor 3.
OS   Oryzias latipes (Medaka fish) (Japanese ricefish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Liu R., Hong Y.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY639947; AAT64912.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   765 AA;  87566 MW;  F5D01408748EC703 CRC64;

  Query Match             85.4%;  Score 998;  DB 2;  Length 765;
  Best Local Similarity   83.5%;  Pred. No. 3.8e-53;
  Matches  192;  Conservative   20;  Mismatches   16;  Indels    2;  Gaps    2;

Qy          2 HPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLN 60
              |||  |||||||:|| :|||:||||||:|||||::|||||||||||||||||:: ||||
Db        130 HPTGTVVTEKQQILEHNLQDIRKRVQDMEQKMKMLENLQDDFDFNYKTLKSQGELNQDLN 189

Qy         61 GNNQ-SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPE 119
              ||:| : ||||| ||||||:||||:|| ||:|: |||:||:|||| |||||||||||| :
Db        190 GNSQAAATRQKMAQLEQMLSALDQLRRQIVTEMGGLLTAMDYVQKNLTDEELADWKRRQQ 249

Qy        120 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 179
              |||||||||||||||| |||||||||| |||||:|||||||||||||:|||| |||:
Db        250 IACIGGPPNICLDRLETWITSLAESQLQIRQQIKKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy        180 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTNKVRLLVKFPEL 229
              ||:|||||||||||||||||||||||||||||||    | ||||||||
Db        310 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTNKVRLLVKFPEL 359


RESULT 10
Q6GUE7
ID   Q6GUE7      PRELIMINARY;      PRT;   785 AA.
AC   Q6GUE7;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Signal transducer and activator of transcription 3 isoform 1.
OS   Oryzias latipes (Medaka fish) (Japanese ricefish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Liu R., Hong Y.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY641434; AAT46364.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
```

```
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   785 AA;  89643 MW;  81F231BDE27DE938 CRC64;

  Query Match             85.4%;  Score 998;  DB 2;  Length 785;
  Best Local Similarity   83.5%;  Pred. No. 3.9e-53;
  Matches  192;  Conservative   20;  Mismatches   16;  Indels    2;  Gaps    2;

Qy          2 HPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLN 60
              |||  |||||||:|| :|||:||||||:|||||::||||||||||||||||:: ||||
Db        130 HPTGTVVTEKQQILEHNLQDIRKRVQDMEQKMKMLENLQDDFDFNYKTLKSQGELNQDLN 189

Qy         61 GNNQ-SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPE 119
              ||:| : ||||| |||||:||||:|| ||:|: |||:||:|||| |||||||||||| :
Db        190 GNSQAAATRQKMAQLEQMLSALDQLRRQIVTEMGGLLTAMDYVQKNLTDEELADWKRRQQ 249

Qy        120 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 179
              |||||||||||||||| |||||||||| |||||||||||||||||||:|||| |||:
Db        250 IACIGGPPNICLDRLETWITSLAESQLQIRQQIKKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy        180 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              ||:|||||||||||||||||||||||||||||||    | ||||||||
Db        310 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTNKVRLLVKFPEL 359


RESULT 11
Q6NV46
ID   Q6NV46      PRELIMINARY;      PRT;   786 AA.
AC   Q6NV46;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Stat3 protein.
GN   Name=stat3;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC068320; AAH68320.1; -.
DR   ZFIN; ZDB-GENE-980526-68; stat3.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   786 AA;  90039 MW;  FC7371D0B0E5447E CRC64;

  Query Match             85.4%;  Score 998;  DB 2;  Length 786;
  Best Local Similarity   83.5%;  Pred. No. 3.9e-53;
  Matches  192;  Conservative   21;  Mismatches   15;  Indels    2;  Gaps    2;

Qy          2 HPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLN 60
              |||  |||||||:|| :|||:||||||:|||||::|||||||||||||||:: ||||
Db        130 HPTGTVVTEKQQILEHNLQDIRKRVQDMEQKMKMLENLQDDFDFNYKTLKSAGELSQDLN 189

Qy         61 GNNQ-SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPE 119
              ||:| : ||||| |||||:||||:|| ||:|:|||||||::||| |||||||||||| :
Db        190 GNSQAAATRQKMSQLEQMLSALDQLRRQIVTEMAGLLSAMDFVQKNLTDEELADWKRRQQ 249

Qy        120 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 179
              |||||||||||||||| |||||||||| ||||:|||||||||||||:|||| |||:
Db        250 IACIGGPPNICLDRLETWITSLAESQLQIRQQIRKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy        180 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              ||:|||||||||||||||||||||||||||||||| | | ||||||||
Db        310 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 359


RESULT 12
O93599
ID   O93599      PRELIMINARY;      PRT;   806 AA.
AC   O93599;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Transcription factor.
GN   Name=stat3;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Oates A.C.;
RL   Thesis (1998), University of Melbourne, Australia.
DR   EMBL; AJ005693; CAA06677.1; -.
DR   HSSP; P42227; 1BG1.
DR   ZFIN; ZDB-GENE-980526-68; stat3.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   806 AA;  92151 MW;  74BC4EA401C3C942 CRC64;

  Query Match             85.4%;  Score 998;  DB 2;  Length 806;
  Best Local Similarity   83.5%;  Pred. No. 4e-53;
```

```
  Matches  192;  Conservative   21;  Mismatches   15;  Indels    2;  Gaps    2;

Qy          2 HPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLN 60
              |||  |||||||:|| :|||:||||||:|||||::||||||||||||||| |:: ||||
Db        130 HPTGTVVTEKQQILEHNLQDIRKRVQDMEQKMKMLENLQDDFDFNYKTLKSAGELSQDLN 189

Qy         61 GNNQ-SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPE 119
              ||:| : ||||| |||||:||||:|| ||:|:|||||||::||| |||||||||||| :
Db        190 GNSQAAATRQKMSQLEQMLSALDQLRRQIVTEMAGLLSAMDFVQKNLTDEELADWKRRQQ 249

Qy        120 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 179
              |||||||||||||||| |||||||||| ||||:||||||||||||||:|||| |||:
Db        250 IACIGGPPNICLDRLETWITSLAESQLQIRQQIRKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy        180 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              ||:|||||||||||||||||||||||||||||||| | | ||||||||
Db        310 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 359


RESULT 13
Q90Y16
ID   Q90Y16      PRELIMINARY;      PRT;   764 AA.
AC   Q90Y16;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   STAT3.
OS   Tetraodon fluviatilis (Puffer fish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=47145;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sung S.C., Fan T.J., Leu J.H., Chou C.M., Huang C.J.;
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF307106; AAL09415.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   764 AA;  87411 MW;  E661FFE18BEFD8BE CRC64;

  Query Match             83.3%;  Score 974;  DB 2;  Length 764;
  Best Local Similarity   82.2%;  Pred. No. 1.1e-51;
  Matches  189;  Conservative   21;  Mismatches   18;  Indels    2;  Gaps    2;

Qy          2 HPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLN 60
              :|:  |||||||||| :|||:||||||:|||||::|||||||||||||||:: ||||
Db        130 NPSGTVVTEKQQMLELNLQDIRKRVQDMEQKMKMLENLQDDFDFNYKTLKSQGELSQDLN 189

Qy         61 GNNQ-SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPE 119
              ||:| |  |||| |||||:||||:|| ||:|: |||:||:|||| ||:|||||||| :
Db        190 GNSQASAARQKMAQLEQMLSALDQLRRQIVTEMGGLLTAMDYVQKNLTNEELADWKRRQQ 249

Qy        120 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 179
              |||||||||||||||| |||||||||| ||||||||||||||||||:|||| |||:
Db        250 IACIGGPPNICLDRLETWITSLAESQLQIRQQIKKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy        180 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              ||:|||||||||||||||||||||||||||||||||    | ||||| ||
Db        310 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTNKVRLLVKFQEL 359
```

```
RESULT 14
O13133
ID   O13133      PRELIMINARY;      PRT;   767 AA.
AC   O13133;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat3.
GN   Name=rbtStat3;
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8022;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Johnson M.C., Mourich D.V., Leong J.C.;
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U60333; AAB60926.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
SQ   SEQUENCE   767 AA;  87816 MW;  FED97740C74C3798 CRC64;

  Query Match             82.5%;  Score 964;  DB 2;  Length 767;
  Best Local Similarity   80.1%;  Pred. No. 4.6e-51;
  Matches  185;  Conservative   25;  Mismatches   19;  Indels    2;  Gaps    2;

Qy          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDL 59
              :||:  |||||||:|| :|||:||||||:|||||::|||||||:|||||||:: ||:
Db        129 SHPSGTVVTEKQQILEHNLQDIRKRVQDMEQKMKMLENLQDDFDFHYKTLKSQGELSQDM 188

Qy         60 NGNNQ-SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRP 118
              |||:| : ||||| |||||:||||:|| ||:|:|||||||::||| |||:||||||||
Db        189 NGNSQAAATRQKMSQLEQMLSALDQLRRQIVTEMAGLLSAMDFVQKNLTDDELADWKRRQ 248

Qy        119 EIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEE 178
              :|||||||| ||||||| ||||| | ||| |||||||||||||||||||:|||| |||
Db        249 QIACIGGPPKICLDRLETWITSLGEIQLQIRQQIKKLEELQQKVSYKGDPIIQHRPALEE 308

Qy        179 RIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVALLVKFPEL 229
              :||:|||||||||||||| ||||||||||||||||    | ||||||||
Db        309 KIVDLFRNLMKSAFVVERQRCMPMHPDRPLVIKTGVQFTNKVRLLVKFPEL 359


RESULT 15
Q9N145
ID   Q9N145      PRELIMINARY;      PRT;   163 AA.
AC   Q9N145;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Signal transducer and activator of transcription (Fragment).
GN   Name=STAT3;
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Arredondo J.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; AF227560; AAF73401.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
FT   NON_TER       1      1
FT   NON_TER     163    163
SQ   SEQUENCE   163 AA;  18855 MW;  DD3F11E7F5B9878C CRC64;

  Query Match             54.7%;  Score 639;  DB 2;  Length 163;
  Best Local Similarity   96.1%;  Pred. No. 6.9e-32;
  Matches  123;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        102 VQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQK 161
              |||||||||||||||| :||||||||||||||||||||||||||||||||||||||||||
Db          1 VQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQK 60

Qy        162 VSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFATAVA 221
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| | ||
Db         61 VSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVR 120

Qy        222 LLVKFPEL 229
              ||||||||
Db        121 LLVKFPEL 128


Search completed: May 25, 2005, 17:43:43
Job time : 96.284 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:14:45 ; Search time 121.891 Seconds
(without alignments)
726.619 Million cell updates/sec

Title: US-10-090-185-30
Perfect score: 1172
Sequence: 1 NHPTAAVVTEKQQMLEQHAQ.....IKTGVQFTTKVRLLVKFPEL 229

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1172	100.0	229	4	AAY72862	Aay72862 Mouse Sta
2	1167	99.6	229	4	AAY72850	Aay72850 Mouse Sta
3	1167	99.6	252	4	AAY72846	Aay72846 Mouse Sta
4	1167	99.6	271	4	AAY72841	Aay72841 Mouse Sta
5	1167	99.6	770	2	AAR72082	Aar72082 Mouse Sta
6	1167	99.6	770	2	AAW03176	Aaw03176 Mouse STA
7	1163	99.2	229	4	AAY72863	Aay72863 Mouse Sta
8	1156	98.6	720	5	AAE22055	Aae22055 Human Sta
9	1156	98.6	769	5	ABB57164	Abb57164 Mouse isc
10	1156	98.6	769	5	AAE22054	Aae22054 Human Sta
11	1156	98.6	769	5	AAE22056	Aae22056 Human pro
12	1156	98.6	770	2	AAR82995	Aar82995 Mouse liv
13	1156	98.6	770	2	AAY03768	Aay03768 Human STA
14	1156	98.6	770	3	AAB12377	Aab12377 N-termina
15	1156	98.6	770	5	AAE14652	Aae14652 Murine ST
16	1156	98.6	770	5	ABG69497	Abg69497 Human bai
17	1156	98.6	770	6	ABU10476	Abu10476 Mouse STA
18	1156	98.6	770	8	ADN04365	Adn04365 Antipsori
19	1156	98.6	770	8	ADP54789	Adp54789 Human PRO
20	1156	98.6	793	3	AAB58442	Aab58442 Lung canc
21	1153	98.4	770	7	ADD44738	Add44738 Rat Prote
22	1152.5	98.3	228	4	AAY72861	Aay72861 Mouse Sta
23	1151	98.2	229	4	AAY72860	Aay72860 Mouse Sta
24	1151	98.2	770	2	AAR82993	Aar82993 Human pla
25	1151	98.2	770	4	AAB19964	Aab19964 Human sig
26	1151	98.2	770	5	AAE15174	Aae15174 Human Sta
27	1151	98.2	770	7	ADD44740	Add44740 Human Pro
28	1089	92.9	213	4	AAY72851	Aay72851 Mouse Sta
29	1089	92.9	236	4	AAY72847	Aay72847 Mouse Sta
30	1044	89.1	223	4	AAY72854	Aay72854 Mouse Sta
31	845	72.1	185	4	AAY72855	Aay72855 Mouse Sta
32	776	66.2	176	4	AAY72848	Aay72848 Mouse Sta
33	653	55.7	128	4	AAY72852	Aay72852 Mouse Sta
34	600	51.2	143	4	AAY72849	Aay72849 Mouse Sta
35	576	49.1	749	5	AAG78526	Aag78526 Rat STAT-
36	572	48.8	268	4	AAY72844	Aay72844 Mouse Sta
37	572	48.8	582	2	AAW62996	Aaw62996 Human tru
38	572	48.8	582	6	ABU04748	Abu04748 Human exp
39	572	48.8	582	8	ADH57036	Adh57036 Truncated
40	572	48.8	680	6	ABR59713	Abr59713 Human sig
41	572	48.8	712	2	AAR72079	Aar72079 Human Sta
42	572	48.8	712	2	AAW03170	Aaw03170 Human STA
43	572	48.8	712	2	AAW62995	Aaw62995 Human Sta
44	572	48.8	712	6	ABU04747	Abu04747 Human exp
45	572	48.8	712	6	ABU04735	Abu04735 Human exp

ALIGNMENTS

RESULT 1
AAY72862
ID AAY72862 standard; protein; 229 AA.
XX
AC AAY72862;
XX
DT 31-MAY-2001 (first entry)
XX
DE Mouse Stat3 mutant (L148A) protein fragment (130-358 amino acids).
XX
KW Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW cellular transformation; dysproliferative disease; cancer; psoriasis;
KW therapy; mutant; mutein.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT Region 1. .25
FT /note= "Stat3-c-Jun interaction region 1; corresponds to
FT 130-154 position of Stat3 protein"
FT Misc-difference 19
FT /note= "Wild type Leu substituted with Ala; corresponds
FT to 148 position of Stat-3 protein"
FT Region 213. .229
FT /note= "Stat3-c-Jun interaction region 2; corresponds to
FT 342-358 position of Stat3 protein"
XX
PN WO200116605-A2.
XX
PD 08-MAR-2001.
XX
PF 30-AUG-2000; 2000WO-US023822.
XX
PR 31-AUG-1999; 99US-00387418.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI Zhang X, Horvath C, Wrzeszcynska MH, Darnell JE;
XX
DR WPI; 2001-226705/23.
XX
PT Identifying an agent for use in modulating the interaction between
PT transcription factor c-Jun and a Stat3 protein.
XX
PS Claim 66; Page 85; 86pp; English.
XX
CC The present sequence is mouse Stat3 mutant (L148A) protein fragment
CC containing 130-358 amino acids of Stat3 protein. This mutant is obtained

```
CC   by replacing Leu 148 with Ala in Stat3 protein. The invention relates to
CC   methods for identifying interacting regions of transcription factors and
CC   methods for identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 229 AA;

  Query Match            100.0%;  Score 1172;  DB 4;  Length 229;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-101;
  Matches  229;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 2
AAY72850
ID   AAY72850 standard; protein; 229 AA.
XX
AC   AAY72850;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 protein fragment #8 (130-358 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .25
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
FT   Region          213. .229
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 76-77; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 130-358
CC   amino acids of Stat3 protein. This Stat3 fragment showed strong binding
CC   to c-Jun protein in the cell extract. The invention relates to methods
CC   for identifying interacting regions of transcription factors and methods
CC   for identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 229 AA;

  Query Match            99.6%;  Score 1167;  DB 4;  Length 229;
  Best Local Similarity  99.6%;  Pred. No. 3.5e-101;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 3
AAY72846
ID   AAY72846 standard; protein; 252 AA.
XX
AC   AAY72846;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 protein fragment #4 (107-358 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          24. .48
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
FT   Region          236. .252
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
```

```
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 73; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 107-358
CC   amino acids of Stat3 protein. This Stat3 fragment showed strong binding
CC   to c-Jun protein in the cell extract. The invention relates to methods
CC   for identifying interacting regions of transcription factors and methods
CC   for identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 252 AA;

  Query Match             99.6%;  Score 1167;  DB 4;  Length 252;
  Best Local Similarity   99.6%;  Pred. No. 4e-101;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db         24 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 83

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 203

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 252


RESULT 4
AAY72841
ID   AAY72841 standard; protein; 271 AA.
XX
AC   AAY72841;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 protein fragment #2 (107-377 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          24. .48
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
FT   Region          236. .252
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 67-68; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 107-377
CC   amino acids of Stat3 protein. This Stat3 fragment showed strong binding
CC   to c-Jun protein in the cell extract. The invention relates to methods
CC   for identifying interacting regions of transcription factors and methods
CC   for identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 271 AA;

  Query Match             99.6%;  Score 1167;  DB 4;  Length 271;
  Best Local Similarity   99.6%;  Pred. No. 4.4e-101;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db         24 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 83

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 203

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 252


RESULT 5
AAR72082
ID   AAR72082 standard; protein; 770 AA.
XX
AC   AAR72082;
XX
DT   25-MAR-2003  (revised)
DT   27-SEP-1995  (first entry)
XX
DE   Mouse Stat3 (19sf6).
XX
KW   Signal transducer and activator of transcription; STAT; 19sf6; Stat3;
KW   receptor recognition factor; transcription factor; cellular debilitation;
KW   derangement; dysfunction; interferon-gamma.
XX
OS   Mus sp.
XX
PN   WO9508629-A1.
XX
PD   30-MAR-1995.
XX
PF   26-SEP-1994;   94WO-US010849.
XX
PR   24-SEP-1993;   93US-00126588.
PR   24-SEP-1993;   93US-00126595.
```

```
PR   11-MAR-1994;   94US-00212184.
PR   11-MAR-1994;   94US-00212185.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Darnell JE,  Schindler CW,  Shuai K,  Wen Z,  Zhong Z;
XX
DR   WPI; 1995-139598/18.
DR   N-PSDB; AAQ89340.
XX
PT   Receptor recognition factor implicated in transcriptional stimulation of
PT   genes - useful in drug screening assays and/or for treating cellular
PT   debilitations, derangements and/or dysfunctions, etc.
XX
PS   Claim 1; Page 107-110; 160pp; English.
XX
CC   A fragment encoding the human Stat91 protein was used to screen a murine
CC   thymus and spleen cDNA for homologous proteins. A highly homologous gene
CC   (given in AAQ89338) was isolated that encoded a 91 kDa protein (AAR72080)
CC   (Stat1) that was responsive to interferon- gamma. Using a fragment of the
CC   mouse gene as probe, 2 additional members of the 113-91 family of
CC   receptor recognition factor proteins were isolated. The 2 genes (AAQ89339
CC   -40) were cloned in plasmids 13sf1 and 19sf6 and encoded proteins termed
CC   Stat4 (AAR72081) and Stat3 (AAR72082), respectively. (Updated on 25-MAR-
CC   2003 to correct PN field.)
XX
SQ   Sequence 770 AA;

  Query Match             99.6%;  Score 1167;  DB 2;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 1.8e-100;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 6
AAW03176
ID   AAW03176 standard; protein; 770 AA.
XX
AC   AAW03176;
XX
DT   24-OCT-1996  (first entry)
XX
DE   Mouse STAT4.
XX
KW   STAT; STAT4; signal transducer and activator of transcription;
KW   DNA binding protein; ligand; receptor; oncogenesis; inflammation;
KW   autoimmune disease; antagonist; therapy.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Domain          398. .508
FT                   /label= DNA_binding_domain
FT                   /note= "Claim 3, page 110"
XX
PN   WO9620954-A2.
XX
PD   11-JUL-1996.
XX
PF   28-DEC-1995;   95WO-US017025.
XX
PR   06-JAN-1995;   95US-00369796.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Darnell JE,  Wen Z,  Horvath CM,  Zhong Z;
XX
DR   WPI; 1996-333941/33.
DR   N-PSDB; AAT31280.
XX
PT   New STAT protein DNA-binding domain peptide(s) - useful for diagnosing,
PT   preventing or treating cellular dysfunction, e.g. oncogenesis,
PT   inflammation, parasitic disease or autoimmunity.
XX
PS   Disclosure; Page 87-90; 138pp; English.
XX
CC   Mouse signal transducer and activator of transcription (STAT) protein
CC   STAT4 (AAW03176) serves a dual purpose, i.e. signal transduction from
CC   ligand-activated receptor kinase complexes followed by nuclear
CC   translocation and DNA binding to activate transcription. Recombinant
CC   STAT4 can be obtd. using cDNA clone 19sf6 (AAT31278) obtd. from
CC   splenic/thymic cells. STAT4 includes a DNA-binding domain (see also
CC   AAW03167) capable of both receptor recognition and message delivery via
CC   DNA binding in a receptor-ligand specific manner. STAT proteins and their
CC   DNA binding domains (see also AAW03165-75) are useful for screening
CC   antagonists used to inhibit STAT-mediated signal transduction and
CC   activation of transcription
XX
SQ   Sequence 770 AA;

  Query Match             99.6%;  Score 1167;  DB 2;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 1.8e-100;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 7
AAY72863
ID   AAY72863 standard; protein; 229 AA.
XX
AC   AAY72863;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 mutant (V151A) protein fragment (130-358 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy; mutant; mutein.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .25
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
```

```
FT   Misc-difference 22
FT                   /note= "Wild type Val substituted with Ala corresponds to
FT                   151 position of Stat-3 protein"
FT   Region          213. .229
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 66; Page 86; 86pp; English.
XX
CC   The present sequence is mouse Stat3 mutant (V151A) protein fragment
CC   containing 130-358 amino acids of Stat3 protein. This mutant is obtained
CC   by replacing Val 151 with Ala in Stat3 protein. The invention relates to
CC   methods for identifying interacting regions of transcription factors and
CC   methods for identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 229 AA;

  Query Match             99.2%;  Score 1163;  DB 4;  Length 229;
  Best Local Similarity   99.1%;  Pred. No. 8.3e-101;
  Matches  227;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||| || ||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 8
AAE22055
ID   AAE22055 standard; protein; 720 AA.
XX
AC   AAE22055;
XX
DT   25-JUL-2002  (first entry)
XX
DE   Human Stat3beta protein.
XX
KW   Human; signal transducer and activator of transcription 3; ischaemia;
KW   immune response; Stat3; coronary atherosclerosis; vascular occlusion;
KW   hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;
KW   inflammation; chronic obstructive pulmonary disease; cardiac arrest;
KW   insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;
KW   shock; chronic active hepatitis; adult respiratory distress syndrome;
KW   nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;
KW   Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;
KW   polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;
KW   proliferative disease; Grave's disease; ulcerative colitis; sarcoma;
KW   carcinoma; degenerative disorder; gene therapy; growth deficiency;
KW   cirrhosis; hypoproliferative disorder; lesion; Statbeta.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 713. .714
FT                   /note= "Encoded by ACA CCA TTC"
XX
PN   WO200220032-A1.
XX
PD   14-MAR-2002.
XX
PF   10-SEP-2001; 2001WO-US028254.
XX
PR   08-SEP-2000; 2000US-0231212P.
XX
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (UYSF-) UNIV SOUTH FLORIDA.
XX
PI   Yu H,  Pardoll D,  Jove R,  Dalton W;
XX
DR   WPI; 2002-362218/39.
DR   N-PSDB; AAD35066.
XX
PT   Modulating angiogenesis and an immune response in an individual, for
PT   treating a hypoxic or ischemic condition, comprises administering a
PT   compound that modulates the activity of a signal transducer and activator
PT   of transcription 3.
XX
PS   Disclosure; Page 87-89; 94pp; English.
XX
CC   The invention relates to a method of modulating angiogenesis and immune
CC   response. Method involves administering to an individual a compound that
CC   modulate the activity of signal transducer and activator of transcription
CC   3 (Stat3). Modulating angiogenesis is useful for treating or preventing
CC   hypoxic or ischaemic condition or disorder which is the result of stroke,
CC   ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,
CC   tissue ischaemia in the lower extremities, infarction, trauma, vascular
CC   occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,
CC   chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,
CC   epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest,
CC   nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy
CC   with neovascularisation. Suppressing an immune response is useful for
CC   ameliorating a symptom of an autoimmune disease such as systemic lupus
CC   erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,
CC   Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,
CC   mixed connective tissue disease, primary biliary cirrhosis, pernicious
CC   anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,
CC   gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,
CC   idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's
CC   disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune
CC   infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and
CC   dense deposit disease. The method is useful in preventing or treating
CC   specific proliferative and oncogenic disease which includes sarcomas and
CC   carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,
CC   fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,
CC   hypoproliferative disorders, physical trauma, lesions and wounds. The
CC   method is also used in gene therapy. The present sequence is human
CC   Stat3beta protein
XX
SQ   Sequence 720 AA;

  Query Match             98.6%;  Score 1156;  DB 5;  Length 720;
  Best Local Similarity   98.7%;  Pred. No. 1.8e-99;
```

```
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 9
ABB57164
ID   ABB57164 standard; protein; 769 AA.
XX
AC   ABB57164;
XX
DT   07-MAR-2002  (first entry)
XX
DE   Mouse ischaemic condition related protein sequence SEQ ID NO:398.
XX
KW   Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW   vasospastic ischaemia; ischaemic condition; ischaemic disease.
XX
OS   Mus musculus.
XX
PN   WO200188188-A2.
XX
PD   22-NOV-2001.
XX
PF   18-MAY-2001; 2001WO-JP004192.
XX
PR   18-MAY-2000; 2000JP-00145977.
XX
PA   (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI   Ishikawa K,  Asai S,  Takahashi Y,  Nagata T,  Ishii Y;
XX
DR   WPI; 2002-034733/04.
DR   N-PSDB; ABI99454.
XX
PT   Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT   expression levels of particular genes defined in the specification or by
PT   determining the expression profile of a gene group comprising these
PT   genes.
XX
PS   Claim 2; Page 1084-1087; 2690pp; English.
XX
CC   The present invention describes a method for examining ischaemic
CC   conditions, comprising measuring the expression levels of particular
CC   genes (I) in a test sample or determining the expression profile of a
CC   gene group in the sample comprising genes selected from (I). The method
CC   is useful for examining the ischaemic condition (e.g. compressive
CC   ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC   expression levels of particular genes (ABI99202 to ABI99912, encoding the
CC   protein sequences in ABB57020 to ABB57374) or by determining the
CC   expression profile of a gene group comprising these genes. The expression
CC   levels or expression profiles produced by these genes are used as an
CC   indicator when screening for ischaemic condition-improving drugs or
CC   therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR
CC   primers for a mouse ischaemic condition related sequence, which are used
CC   in the exemplification of the present invention
XX
SQ   Sequence 769 AA;

  Query Match           98.6%;  Score 1156;  DB 5;  Length 769;
  Best Local Similarity   98.7%;  Pred. No. 1.9e-99;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 10
AAE22054
ID   AAE22054 standard; protein; 769 AA.
XX
AC   AAE22054;
XX
DT   25-JUL-2002  (first entry)
XX
DE   Human Stat3 protein.
XX
KW   Human; signal transducer and activator of transcription 3; ischaemia;
KW   immune response; Stat3; coronary atherosclerosis; vascular occlusion;
KW   hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;
KW   inflammation; chronic obstructive pulmonary disease; cardiac arrest;
KW   insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;
KW   shock; chronic active hepatitis; adult respiratory distress syndrome;
KW   nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;
KW   Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;
KW   polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;
KW   proliferative disease; Grave's disease; ulcerative colitis; sarcoma;
KW   carcinoma; degenerative disorder; gene therapy; growth deficiency;
KW   cirrhosis; hypoproliferative disorder; lesion.
XX
OS   Homo sapiens.
XX
PN   WO200220032-A1.
XX
PD   14-MAR-2002.
XX
PF   10-SEP-2001; 2001WO-US028254.
XX
PR   08-SEP-2000; 2000US-0231212P.
XX
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (UYSF-) UNIV SOUTH FLORIDA.
XX
PI   Yu H,  Pardoll D,  Jove R,  Dalton W;
XX
DR   WPI; 2002-362218/39.
DR   N-PSDB; AAD35065.
XX
PT   Modulating angiogenesis and an immune response in an individual, for
PT   treating a hypoxic or ischemic condition, comprises administering a
PT   compound that modulates the activity of a signal transducer and activator
PT   of transcription 3.
XX
PS   Disclosure; Page 83-85; 94pp; English.
XX
CC   The invention relates to a method of modulating angiogenesis and immune
CC   response. Method involves administering to an individual a compound that
CC   modulate the activity of signal transducer and activator of transcription
CC   3 (Stat3). Modulating angiogenesis is useful for treating or preventing
```

CC hypoxic or ischaemic condition or disorder which is the result of stroke,
CC ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,
CC tissue ischaemia in the lower extremities, infarction, trauma, vascular
CC occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,
CC chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,
CC epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest,
CC nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy
CC with neovascularisation. Suppressing an immune response is useful for
CC ameliorating a symptom of an autoimmune disease such as systemic lupus
CC erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,
CC Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,
CC mixed connective tissue disease, primary biliary cirrhosis, pernicious
CC anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,
CC gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,
CC idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's
CC disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune
CC infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and
CC dense deposit disease. The method is useful in preventing or treating
CC specific proliferative and oncogenic disease which includes sarcomas and
CC carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,
CC fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,
CC hypoproliferative disorders, physical trauma, lesions and wounds. The
CC method is also used in gene therapy. The present sequence is human Stat3
CC protein
XX
SQ Sequence 769 AA;

Query Match 98.6%; Score 1156; DB 5; Length 769;
Best Local Similarity 98.7%; Pred. No. 1.9e-99;
Matches 226; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358
```

RESULT 11
AAE22056
ID AAE22056 standard; protein; 769 AA.
XX
AC AAE22056;
XX
DT 25-JUL-2002 (first entry)
XX
DE Human protein related to angiogenesis regulation.
XX
KW Human; signal transducer and activator of transcription 3; ischaemia;
KW immune response; Stat3; coronary atherosclerosis; vascular occlusion;
KW hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;
KW inflammation; chronic obstructive pulmonary disease; cardiac arrest;
KW insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;
KW shock; chronic active hepatitis; adult respiratory distress syndrome;
KW nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;
KW Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;
KW polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;
KW proliferative disease; Grave's disease; ulcerative colitis; sarcoma;
KW carcinoma; degenerative disorder; gene therapy; growth deficiency;
KW cirrhosis; hypoproliferative disorder; lesion.
XX
OS Homo sapiens.
XX
PN WO200220032-A1.
XX
PD 14-MAR-2002.
XX
PF 10-SEP-2001; 2001WO-US028254.
XX
PR 08-SEP-2000; 2000US-0231212P.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
PA (UYSF-) UNIV SOUTH FLORIDA.
XX
PI Yu H, Pardoll D, Jove R, Dalton W;
XX
DR WPI; 2002-362218/39.
XX
PT Modulating angiogenesis and an immune response in an individual, for
PT treating a hypoxic or ischemic condition, comprises administering a
PT compound that modulates the activity of a signal transducer and activator
PT of transcription 3.
XX
PS Disclosure; Page 83-85; 94pp; English.
XX
CC The invention relates to a method of modulating angiogenesis and immune
CC response. Method involves administering to an individual a compound that
CC modulate the activity of signal transducer and activator of transcription
CC 3 (Stat3). Modulating angiogenesis is useful for treating or preventing
CC hypoxic or ischaemic condition or disorder which is the result of stroke,
CC ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,
CC tissue ischaemia in the lower extremities, infarction, trauma, vascular
CC occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,
CC chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,
CC epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest,
CC nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy
CC with neovascularisation. Suppressing an immune response is useful for
CC ameliorating a symptom of an autoimmune disease such as systemic lupus
CC erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,
CC Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,
CC mixed connective tissue disease, primary biliary cirrhosis, pernicious
CC anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,
CC gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,
CC idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's
CC disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune
CC infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and
CC dense deposit disease. The method is useful in preventing or treating
CC specific proliferative and oncogenic disease which includes sarcomas and
CC carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,
CC fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,
CC hypoproliferative disorders, physical trauma, lesions and wounds. The
CC method is also used in gene therapy. The present sequence is human
CC protein related to angiogenesis regulation
XX
SQ Sequence 769 AA;

Query Match 98.6%; Score 1156; DB 5; Length 769;
Best Local Similarity 98.7%; Pred. No. 1.9e-99;
Matches 226; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358
```

```
RESULT 12
AAR82995
ID   AAR82995 standard; protein; 770 AA.
XX
AC   AAR82995;
XX
DT   25-MAR-1996  (first entry)
XX
DE   Mouse liver acute phase response factor.
XX
KW   Mouse; acute phase response factor; transcription factor; interleukin-6;
KW   signal transmission; liver; antibody; antisense; ribozyme;
KW   antiinflammatory; antitumor; hypotensive; therapy.
XX
OS   Mus musculus.
XX
PN   EP676469-A2.
XX
PD   11-OCT-1995.
XX
PF   29-MAR-1995;   95EP-00104670.
XX
PR   04-APR-1994;   94JP-00065825.
XX
PA   (KISH/) KISHIMOTO T.
XX
PI   Akira S,  Kishimoto T;
XX
DR   WPI; 1995-346089/45.
DR   N-PSDB; AAT05619.
XX
PT   New acute phase response factor - for developing inhibitory agents for
PT   treating diseases induced by cytokine(s) such as IL-6, e.g. inflammatory
PT   diseases.
XX
PS   Claim 10; Page 20-22; 31pp; English.
XX
CC   The sequence represents a mouse acute phase response factor (APRF), a
CC   transcription factor related to signal transmission of interleukin-6 (IL-
CC   6). The protein is encoded by a cDNA, isolated from a mouse liver cDNA
CC   library using a polymerase chain reaction product (amplified using
CC   primers derived from an IL-6-treated mouse liver peptide) as a probe.
CC   APRF-inhibitors, e.g. antibodies, antisense oligonucleotides or
CC   ribozymes, may be used to treat diseases induced by IL-6, e.g.
CC   inflammatory disease, leukemia, cancer, osteoclasia, pulmonary
CC   hypertension, etc
XX
SQ   Sequence 770 AA;

  Query Match             98.6%;  Score 1156;  DB 2;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 1.9e-99;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 13
AAY03768
ID   AAY03768 standard; protein; 770 AA.
XX
AC   AAY03768;
XX
DT   11-JUN-1999  (first entry)
XX
DE   Human STAT3 allelic variant.
XX
KW   Signal Transducer and Activator of Transcription 3; STAT3; allele; IL-6;
KW   intracellular transcription factor; interleukin-6; medicament; variant;
KW   pharmaceutical; autoimmune disease; inflammatory; human.
XX
OS   Homo sapiens.
XX
PN   EP905234-A2.
XX
PD   31-MAR-1999.
XX
PF   18-FEB-1998;   98EP-00102774.
XX
PR   16-SEP-1997;   97EP-00116061.
XX
PA   (ISTF ) ARS APPLIED RES SYSTEMS HOLDING NV.
XX
PI   Serlupi-Crescenzi O,  Della Pietra L;
XX
DR   WPI; 1999-192664/17.
DR   N-PSDB; AAX29281.
XX
PT   New human Signal Transducer and Activator of Transcription 3 (STAT3)
PT   allelic variant useful for treatment of autoimmune and inflammatory
PT   disease.
XX
PS   Claim 2; Page 9-13; 32pp; English.
XX
CC   The present sequence represents a predominant allelic variant of human
CC   Signal Transducer and Activator of Transcription 3 (STAT3) protein, an
CC   intracellular transcription factor which mediates IL-6 signals. The
CC   encoding sequence differs from the original published human STAT3 gene
CC   sequence (Akira et al., 1994). Host cells comprising a recombinant STAT3
CC   DNA molecule can be used for the recombinant expression of the variant.
CC   STAT3 protein is useful as a medicament or pharmaceutical composition for
CC   treatment of autoimmune or inflammatory diseases
XX
SQ   Sequence 770 AA;

  Query Match             98.6%;  Score 1156;  DB 2;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 1.9e-99;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 14
AAB12377
ID   AAB12377 standard; peptide; 770 AA.
XX
AC   AAB12377;
XX
```

```
DT   08-NOV-2000  (first entry)
XX
DE   N-terminal domain of murine STAT-3 protein.
XX
KW   STAT; signal transducer and activator of transcription; crystal;
KW   drug design; murine.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Region          4. .9
FT                   /label= Alpha helix 1
FT   Region          12. .21
FT                   /label= Alpha helix 2
FT   Region          19. .21
FT                   /label= 3(10) helix of alpha helix 2
FT   Region          28. .33
FT                   /label= Alpha helix 3
FT   Region          35. .40
FT                   /label= Alpha helix 4
FT   Region          43. .47
FT                   /label= Alpha helix 5
FT   Region          50. .73
FT                   /label= Alpha helix 6
FT   Region          77. .96
FT                   /label= Alpha helix 7
FT   Region          99. .119
FT                   /label= Alpha helix 8
XX
PN   US6087478-A.
XX
PD   11-JUL-2000.
XX
PF   23-JAN-1998;   98US-00012710.
XX
PR   23-JAN-1998;   98US-00012710.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Vinkemeier U,  Moarefi I,  Darnell JE,  Kuriyan J;
XX
DR   WPI; 2000-505108/45.
XX
PT   New crystals of an N-terminal fragment of a signal transducer and
PT   activator of transcription that effectively diffracts x-rays, useful for
PT   drug screening and development.
XX
PS   Disclosure; Fig 1; 42pp; English.
XX
CC   The present invention relates to a crystal of an N-terminal fragment of a
CC   signal transducer and activator of transcription (STAT) protein. The
CC   crystal effectively diffracts X-rays, allowing the determination of the
CC   atomic coordinates of the N-terminal domain to a resolution of greater
CC   than 5.0 Angstroms. The present sequence is the N-terminal domain of the
CC   murine STAT 3 protein. The N-terminal domain enables STAT dimers to
CC   interact and bind DNA cooperatively, a mechanism important for gene
CC   activation. The crystals are useful in drug screening and development by
CC   selecting a potential drug by performing rational drug design with the 3-
CC   dimensional structure determined for the crystal
XX
SQ   Sequence 770 AA;

  Query Match             98.6%;  Score 1156;  DB 3;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 1.9e-99;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 15
AAE14652
ID   AAE14652 standard; protein; 770 AA.
XX
AC   AAE14652;
XX
DT   16-JUL-2002  (first entry)
XX
DE   Murine STAT3 protein.
XX
KW   Signal transducer and activator of transcription; STAT3;
KW   drug development; drug discovery; crystal; inflammation; allergy; asthma;
KW   leukaemia; anaemia; neutropaenia; thrombocytopaenia; cancer; obesity;
KW   viral disease; growth retardation; murine.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Domain          1. .130
FT                   /note= "Conserved N-terminal domain of the STAT family"
XX
PN   US6312887-B1.
XX
PD   06-NOV-2001.
XX
PF   24-APR-2000; 2000US-00556273.
XX
PR   23-JAN-1998;   98US-00012710.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Vinkemeier U,  Moarefi I,  Darnell JE,  Kuriyan J;
XX
DR   WPI; 2002-033337/04.
XX
PT   Identifying compounds that bind to signal transducer and activator of
PT   transcription proteins, useful for the production of new drugs.
XX
PS   Example; Col 47-50; 44pp; English.
XX
CC   The invention relates to methods for detecting compounds that bind to
CC   signal transducer and activator of transcription (STAT) proteins for the
CC   discovery and development of new drug compounds based on the structural
CC   properties of the protein crystal. The methods include: identifying a
CC   compound that binds to the N-terminal domain of a STAT protein,
CC   identifying a compound that enhances or diminishes the binding of the
CC   dimeric STAT proteins to each other and/or their nucleic acid binding
CC   site; or identifying a compound that enhances or diminishes the ability
CC   of STAT protein dimers to induce the expression of a gene operably under
CC   the control of a promoter containing at least two adjacent weak binding
CC   sites for STAT protein dimers. The methods are used for identifying new
CC   drugs. An antagonist of STAT N-terminal dimeric interactions that
CC   inhibits the binding of the STAT dimers to adjacent weak binding sites on
CC   a promoter of a gene, could be useful as drugs in the treatment of
CC   diseases, e.g. inflammation, allergy, asthma and leukaemias. On the other
CC   hand, an agonist of N-terminal dimeric interactions between STAT dimers,
CC   can be used as drugs in the treatment of diseases e.g. anaemia,
CC   neutropaenia, thrombocytopaenia, cancer, obesity, viral diseases and
CC   growth retardation. The present sequence is murine STAT3 protein
XX
SQ   Sequence 770 AA;

  Query Match             98.6%;  Score 1156;  DB 5;  Length 770;
```

Best Local Similarity 98.7%; Pred. No. 1.9e-99;
Matches 226; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358
```

Search completed: May 25, 2005, 17:36:38
Job time : 123.057 secs

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:25:01 ; Search time 28.6534 Seconds
(without alignments)
596.600 Million cell updates/sec

Title: US-10-090-185-30
Perfect score: 1172
Sequence: 1 NHPTAAVVTEKQQMLEQHAQ.....IKTGVQFTTKVRLLVKFPEL 229

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1172	100.0	229	3	US-09-387-418A-30	Sequence 30, Appl
2	1167	99.6	229	3	US-09-387-418A-18	Sequence 18, Appl
3	1167	99.6	229	3	US-09-387-418A-28	Sequence 28, Appl
4	1167	99.6	252	3	US-09-387-418A-14	Sequence 14, Appl
5	1167	99.6	271	3	US-09-387-418A-9	Sequence 9, Appli
6	1167	99.6	770	1	US-08-369-796-12	Sequence 12, Appl
7	1167	99.6	770	2	US-08-852-091-12	Sequence 12, Appl
8	1167	99.6	770	2	US-08-820-754-12	Sequence 12, Appl
9	1167	99.6	770	3	US-08-956-652-12	Sequence 12, Appl
10	1167	99.6	770	3	US-08-956-869-12	Sequence 12, Appl
11	1167	99.6	770	3	US-08-948-547-12	Sequence 12, Appl
12	1167	99.6	770	3	US-09-364-970-3	Sequence 3, Appli
13	1167	99.6	770	3	US-09-364-970-5	Sequence 5, Appli
14	1167	99.6	770	3	US-08-956-653A-12	Sequence 12, Appl
15	1167	99.6	770	4	US-08-212-185-12	Sequence 12, Appl
16	1167	99.6	770	5	PCT-US95-17025-12	Sequence 12, Appl
17	1163	99.2	229	3	US-09-387-418A-31	Sequence 31, Appl
18	1156	98.6	770	1	US-08-416-581B-9	Sequence 9, Appli
19	1156	98.6	770	3	US-09-012-710-8	Sequence 8, Appli
20	1156	98.6	770	3	US-09-556-273-8	Sequence 8, Appli
21	1156	98.6	770	3	US-09-526-542-2	Sequence 2, Appli
22	1156	98.6	770	4	US-10-117-087-2	Sequence 2, Appli
23	1151	98.2	770	1	US-08-416-581B-1	Sequence 1, Appli
24	1151	98.2	770	1	US-08-416-581B-5	Sequence 5, Appli
25	1151	98.2	770	3	US-09-087-465-6	Sequence 6, Appli
26	1151	98.2	770	4	US-09-972-800A-6	Sequence 6, Appli
27	1151	98.2	771	1	US-08-276-099A-14	Sequence 14, Appl
28	1151	98.2	771	1	US-08-781-890-14	Sequence 14, Appl
29	1150	98.1	229	3	US-09-387-418A-29	Sequence 29, Appl
30	1089	92.9	213	3	US-09-387-418A-19	Sequence 19, Appl
31	1089	92.9	236	3	US-09-387-418A-15	Sequence 15, Appl
32	1044	89.1	223	3	US-09-387-418A-22	Sequence 22, Appl
33	845	72.1	185	3	US-09-387-418A-23	Sequence 23, Appl
34	776	66.2	176	3	US-09-387-418A-16	Sequence 16, Appl
35	653	55.7	128	3	US-09-387-418A-20	Sequence 20, Appl
36	600	51.2	143	3	US-09-387-418A-17	Sequence 17, Appl
37	572	48.8	268	3	US-09-387-418A-12	Sequence 12, Appl
38	572	48.8	582	4	US-09-430-806A-3	Sequence 3, Appli
39	572	48.8	712	1	US-08-369-796-6	Sequence 6, Appli
40	572	48.8	712	2	US-08-852-091-6	Sequence 6, Appli
41	572	48.8	712	2	US-08-820-754-6	Sequence 6, Appli
42	572	48.8	712	3	US-08-956-652-6	Sequence 6, Appli
43	572	48.8	712	3	US-08-956-869-6	Sequence 6, Appli
44	572	48.8	712	3	US-08-948-547-6	Sequence 6, Appli
45	572	48.8	712	3	US-08-956-653A-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-09-387-418A-30
; Sequence 30, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/09/387,418A
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-387-418A-30

Query Match 100.0%; Score 1172; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 1e-97;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
```

RESULT 2
US-09-387-418A-18
; Sequence 18, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H

```
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 18
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-18

  Query Match            99.6%;  Score 1167;  DB 3;  Length 229;
  Best Local Similarity  99.6%;  Pred. No. 2.9e-97;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 3
US-09-387-418A-28
; Sequence 28, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
;  APPLICANT:  Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 28
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-28

  Query Match            99.6%;  Score 1167;  DB 3;  Length 229;
  Best Local Similarity  99.6%;  Pred. No. 2.9e-97;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 4
US-09-387-418A-14
; Sequence 14, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
;  APPLICANT:  Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 252
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-14

  Query Match            99.6%;  Score 1167;  DB 3;  Length 252;
  Best Local Similarity  99.6%;  Pred. No. 3.3e-97;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db         24 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 83

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 203

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 252


RESULT 5
US-09-387-418A-9
; Sequence 9, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
;  APPLICANT:  Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 271
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-9

  Query Match            99.6%;  Score 1167;  DB 3;  Length 271;
  Best Local Similarity  99.6%;  Pred. No. 3.6e-97;
```

```
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db         24 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 83

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 203

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 252


RESULT 6
US-08-369-796-12
; Sequence 12, Application US/08369796
; Patent No. 5716622
;  GENERAL INFORMATION:
;    APPLICANT:  James E. Darnell, Jr.
;    APPLICANT:  Zilong Wen
;    APPLICANT:  Curt M. Horvath
;    APPLICANT:  Zhong Zhong
;    TITLE OF INVENTION:  FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
;    TITLE OF INVENTION:  TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/369,796
;      FILING DATE:  06-JAN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-116
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-369-796-12

  Query Match             99.6%;  Score 1167;  DB 1;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 1.4e-96;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 7
US-08-852-091-12
; Sequence 12, Application US/08852091
; Patent No. 5883228
;  GENERAL INFORMATION:
;    APPLICANT:  James E. Darnell, Jr.
;    APPLICANT:  Zilong Wen
;    APPLICANT:  Curt M. Horvath
;    APPLICANT:  Zhong Zhong
;    TITLE OF INVENTION:  FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
;    TITLE OF INVENTION:  TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/852,091
;      FILING DATE:  06-MAY-1997
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/369,796
;      FILING DATE:  06-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-116
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-852-091-12

  Query Match             99.6%;  Score 1167;  DB 2;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 1.4e-96;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309
```

```
Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 8
US-08-820-754-12
; Sequence 12, Application US/08820754
; Patent No. 5976835
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/820,754
;      FILING DATE:  19-MAR-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,185
;      FILING DATE:  11-MAR-1994
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-820-754-12

  Query Match             99.6%;  Score 1167;  DB 2;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 1.4e-96;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 9
US-08-956-652-12
; Sequence 12, Application US/08956652
; Patent No. 6013475
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/956,652
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,185
;      FILING DATE:  11-MAR-1994
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-956-652-12

  Query Match             99.6%;  Score 1167;  DB 3;  Length 770;
```

```
  Best Local Similarity   99.6%;  Pred. No. 1.4e-96;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 10
US-08-956-869-12
; Sequence 12, Application US/08956869
; Patent No. 6030808
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/956,869
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/212,185
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-956-869-12

  Query Match             99.6%;  Score 1167;  DB 3;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 1.4e-96;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 11
US-08-948-547-12
; Sequence 12, Application US/08948547
; Patent No. 6124118
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/948,547
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/212,185
;      FILING DATE:  11-MAR-1994
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
```

```
;       TELEPHONE:  201 487-5800
;       TELEFAX:  201 343-1684
;       TELEX:  133521
;   INFORMATION FOR SEQ ID NO:  12:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  770 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-948-547-12

  Query Match             99.6%;  Score 1167;  DB 3;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 1.4e-96;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 12
US-09-364-970-3
; Sequence 3, Application US/09364970
; Patent No. 6235873
; GENERAL INFORMATION:
;  APPLICANT: Bromberg, Jacqueline
;  TITLE OF INVENTION: CONSTITUTIVELY ACTIVE TRANSCRIPTION FACTORS AND THEIR
;  TITLE OF INVENTION: USES FOR IDENTIFYING MODULATORS OF ACTIVITY INCLUDING
;  TITLE OF INVENTION: DYSPROLIFERATIVE CELLULAR CHANGES
;  FILE REFERENCE: 600-1-252
;  CURRENT APPLICATION NUMBER: US/09/364,970
;  CURRENT FILING DATE: 1999-07-31
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-364-970-3

  Query Match             99.6%;  Score 1167;  DB 3;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 1.4e-96;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 13
US-09-364-970-5
; Sequence 5, Application US/09364970
; Patent No. 6235873
; GENERAL INFORMATION:
;  APPLICANT: Bromberg, Jacqueline
;  TITLE OF INVENTION: CONSTITUTIVELY ACTIVE TRANSCRIPTION FACTORS AND THEIR
;  TITLE OF INVENTION: USES FOR IDENTIFYING MODULATORS OF ACTIVITY INCLUDING
;  TITLE OF INVENTION: DYSPROLIFERATIVE CELLULAR CHANGES
;  FILE REFERENCE: 600-1-252
;  CURRENT APPLICATION NUMBER: US/09/364,970
;  CURRENT FILING DATE: 1999-07-31
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-364-970-5

  Query Match             99.6%;  Score 1167;  DB 3;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 1.4e-96;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 14
US-08-956-653A-12
; Sequence 12, Application US/08956653A
; Patent No. 6338949
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/956,653A
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,185
;      FILING DATE:  11-MAR-1994
```

```
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-195
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-956-653A-12

  Query Match             99.6%;  Score 1167;  DB 3;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 1.4e-96;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 15
US-08-212-185-12
; Sequence 12, Application US/08212185
; Patent No. 6605442
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/212,185
;      FILING DATE:  11-MAR-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-212-185-12

  Query Match             99.6%;  Score 1167;  DB 4;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 1.4e-96;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


Search completed: May 25, 2005, 17:47:40
Job time : 29.6534 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:36:58 ; Search time 100.287 Seconds
(without alignments)
763.830 Million cell updates/sec

Title: US-10-090-185-30
Perfect score: 1172
Sequence: 1 NHPTAAVVTEKQQMLEQHAQ.....IKTGVQFTTKVRLLVKFPEL 229

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1434725 seqs, 334507595 residues

Total number of hits satisfying chosen parameters: 1434725

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1172	100.0	229	13	US-10-090-185-30	Sequence 30, Appl
2	1167	99.6	229	13	US-10-090-185-18	Sequence 18, Appl
3	1167	99.6	229	13	US-10-090-185-28	Sequence 28, Appl
4	1167	99.6	252	13	US-10-090-185-14	Sequence 14, Appl
5	1167	99.6	271	13	US-10-090-185-9	Sequence 9, Appli
6	1167	99.6	770	11	US-09-876-773-12	Sequence 12, Appl
7	1167	99.6	770	17	US-10-639-617-12	Sequence 12, Appl
8	1163	99.2	229	13	US-10-090-185-31	Sequence 31, Appl
9	1156	98.6	720	15	US-10-380-020-4	Sequence 4, Appli
10	1156	98.6	769	15	US-10-380-020-2	Sequence 2, Appli
11	1156	98.6	769	15	US-10-380-020-5	Sequence 5, Appli
12	1156	98.6	770	14	US-10-045-792-8	Sequence 8, Appli
13	1156	98.6	770	14	US-10-038-010-56	Sequence 56, Appl
14	1156	98.6	770	14	US-10-117-087-2	Sequence 2, Appli
15	1156	98.6	793	9	US-09-925-302-780	Sequence 780, App
16	1156	98.6	793	10	US-09-925-302-780	Sequence 780, App
17	1153	98.4	770	15	US-10-116-275-329	Sequence 329, App
18	1151	98.2	770	15	US-10-116-275-349	Sequence 349, App
19	1150	98.1	229	13	US-10-090-185-29	Sequence 29, Appl
20	1089	92.9	213	13	US-10-090-185-19	Sequence 19, Appl
21	1089	92.9	236	13	US-10-090-185-15	Sequence 15, Appl
22	1044	89.1	223	13	US-10-090-185-22	Sequence 22, Appl
23	845	72.1	185	13	US-10-090-185-23	Sequence 23, Appl
24	776	66.2	176	13	US-10-090-185-16	Sequence 16, Appl
25	653	55.7	128	13	US-10-090-185-20	Sequence 20, Appl
26	600	51.2	143	13	US-10-090-185-17	Sequence 17, Appl
27	576	49.1	749	9	US-09-833-205-4	Sequence 4, Appli
28	572	48.8	268	13	US-10-090-185-12	Sequence 12, Appl
29	572	48.8	582	14	US-10-245-120-3	Sequence 3, Appli
30	572	48.8	712	11	US-09-876-773-6	Sequence 6, Appli
31	572	48.8	712	14	US-10-245-120-2	Sequence 2, Appli
32	572	48.8	712	17	US-10-639-617-6	Sequence 6, Appli
33	572	48.8	712	17	US-10-936-390-5	Sequence 5, Appli
34	572	48.8	750	9	US-09-833-205-2	Sequence 2, Appli
35	572	48.8	750	11	US-09-876-773-4	Sequence 4, Appli
36	572	48.8	750	14	US-10-245-120-1	Sequence 1, Appli
37	572	48.8	750	14	US-10-308-279-44	Sequence 44, Appl
38	572	48.8	750	16	US-10-755-889-352	Sequence 352, App
39	572	48.8	750	16	US-10-755-889-823	Sequence 823, App
40	572	48.8	750	17	US-10-492-043-19	Sequence 19, Appl
41	572	48.8	750	17	US-10-639-617-4	Sequence 4, Appli
42	572	48.8	786	9	US-09-925-297-550	Sequence 550, App
43	571	48.7	129	13	US-10-090-185-24	Sequence 24, Appl
44	565	48.2	749	9	US-09-833-205-6	Sequence 6, Appli
45	565	48.2	749	11	US-09-876-773-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-10-090-185-30
; Sequence 30, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/10/090,185
; CURRENT FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: 09/387,418
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-090-185-30

Query Match 100.0%; Score 1172; DB 13; Length 229;
Best Local Similarity 100.0%; Pred. No. 2.3e-95;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
```

```
Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 2
US-10-090-185-18
; Sequence 18, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-18

  Query Match             99.6%;  Score 1167;  DB 13;  Length 229;
  Best Local Similarity   99.6%;  Pred. No. 6.4e-95;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 3
US-10-090-185-28
; Sequence 28, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-28

  Query Match             99.6%;  Score 1167;  DB 13;  Length 229;
  Best Local Similarity   99.6%;  Pred. No. 6.4e-95;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 4
US-10-090-185-14
; Sequence 14, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 252
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-14

  Query Match             99.6%;  Score 1167;  DB 13;  Length 252;
  Best Local Similarity   99.6%;  Pred. No. 7.2e-95;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db         24 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 83

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 203

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 252


RESULT 5
US-10-090-185-9
```

```
; Sequence 9, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 271
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-9

  Query Match            99.6%;  Score 1167;  DB 13;  Length 271;
  Best Local Similarity  99.6%;  Pred. No. 7.9e-95;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         24 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 83

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 203

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 252


RESULT 6
US-09-876-773-12
; Sequence 12, Application US/09876773
; Publication No. US20040058318A1
;  GENERAL INFORMATION:
;       APPLICANT: Darnell Jr., James E.
;                  Schindler, Christian W.
;                  Fu, Xian-Yuan
;                  Wen, Zilong
;                  Zhong, Zhong
;       TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
;                           SEQUENCES AND METHODS OF USE THEREOF
;       NUMBER OF SEQUENCES: 25
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Klauber & Jackson
;            STREET: 411 Hackensack Avenue
;            CITY: Hackensack
;            STATE: New Jersey
;            COUNTRY: USA
;            ZIP: 07601
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/876,773
;            FILING DATE: 07-Jun-2001
;            CLASSIFICATION: 435
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/212,185
;            FILING DATE: 11-MAR-1994
;            APPLICATION NUMBER: US 07/980,498
;            FILING DATE: 23-NOV-1992
;            APPLICATION NUMBER: US 07/854,296
;            FILING DATE: 19-MAR-1992
;            APPLICATION NUMBER: WO US93/02569
;            FILING DATE: 19-MAR-1993
;            APPLICATION NUMBER: US 08/126,588
;            FILING DATE: 24-SEP-1993
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Jackson Esq., David A.
;            REGISTRATION NUMBER: 26,742
;            REFERENCE/DOCKET NUMBER: 600-1-073 CIP
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 201 487-5800
;            TELEFAX: 201 343-1684
;            TELEX: 133521
;  INFORMATION FOR SEQ ID NO: 12:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 770 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-876-773-12

  Query Match            99.6%;  Score 1167;  DB 11;  Length 770;
  Best Local Similarity  99.6%;  Pred. No. 3e-94;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 7
US-10-639-617-12
; Sequence 12, Application US/10639617
; Publication No. US20050079543A1
;  GENERAL INFORMATION:
;       APPLICANT: Darnell Jr., James E.
;                  Schindler, Christian W.
;                  Fu, Xian-Yuan
;                  Wen, Zilong
;                  Zhong, Zhong
;       TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
;                           SEQUENCES AND METHODS OF USE THEREOF
;       NUMBER OF SEQUENCES: 25
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Klauber & Jackson
;            STREET: 411 Hackensack Avenue
;            CITY: Hackensack
;            STATE: New Jersey
;            COUNTRY: USA
;            ZIP: 07601
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/639,617
;            FILING DATE: 12-Aug-2003
;            CLASSIFICATION: 435
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/212,185
;            FILING DATE: 11-MAR-1994
;            APPLICATION NUMBER: US 07/980,498
;            FILING DATE: 23-NOV-1992
;            APPLICATION NUMBER: US 07/854,296
;            FILING DATE: 19-MAR-1992
;            APPLICATION NUMBER: WO US93/02569
;            FILING DATE: 19-MAR-1993
;            APPLICATION NUMBER: US 08/126,588
;            FILING DATE: 24-SEP-1993
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Jackson Esq., David A.
;            REGISTRATION NUMBER: 26,742
;            REFERENCE/DOCKET NUMBER: 600-1-073 CIP
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 201 487-5800
;            TELEFAX: 201 343-1684
;            TELEX: 133521
;   INFORMATION FOR SEQ ID NO: 12:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 770 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-639-617-12

  Query Match             99.6%;  Score 1167;  DB 17;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 3e-94;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 8
US-10-090-185-31
; Sequence 31, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-31

  Query Match             99.2%;  Score 1163;  DB 13;  Length 229;
  Best Local Similarity   99.1%;  Pred. No. 1.4e-94;
  Matches  227;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||| || |||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 9
US-10-380-020-4
; Sequence 4, Application US/10380020
; Publication No. US20040052762A1
; GENERAL INFORMATION:
;  APPLICANT; Yu, Hua
;  APPLICANT:  Pardoll, Drew
;  APPLICANT:  Jove, Richard
;  APPLICANT:  Dalton, Willian
;  TITLE OF INVENTION: Stat3 Agonists and Antagonists and Therapeutic Uses Thereof
;  FILE REFERENCE: 10873-009-999
;  CURRENT APPLICATION NUMBER: US/10/380,020
;  CURRENT FILING DATE:  2003-03-07
;  PRIOR APPLICATION NUMBER: 60/231,212
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
;   LENGTH: 720
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-380-020-4

  Query Match             98.6%;  Score 1156;  DB 15;  Length 720;
  Best Local Similarity   98.7%;  Pred. No. 2.6e-93;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 10
US-10-380-020-2
; Sequence 2, Application US/10380020
; Publication No. US20040052762A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Yu, Hua
;  APPLICANT:  Pardoll, Drew
;  APPLICANT:  Jove, Richard
;  APPLICANT:  Dalton, Willian
;  TITLE OF INVENTION: Stat3 Agonists and Antagonists and Therapeutic Uses Thereof
;  FILE REFERENCE: 10873-009-999
;  CURRENT APPLICATION NUMBER: US/10/380,020
;  CURRENT FILING DATE:  2003-03-07
;  PRIOR APPLICATION NUMBER: 60/231,212
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 769
;   TYPE: PRT<
;   ORGANISM: Homo Sapiens
US-10-380-020-2

  Query Match             98.6%;  Score 1156;  DB 15;  Length 769;
  Best Local Similarity   98.7%;  Pred. No. 2.8e-93;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 11
US-10-380-020-5
; Sequence 5, Application US/10380020
; Publication No. US20040052762A1
; GENERAL INFORMATION:
;  APPLICANT: Yu, Hua
;  APPLICANT:  Pardoll, Drew
;  APPLICANT:  Jove, Richard
;  APPLICANT:  Dalton, Willian
;  TITLE OF INVENTION: Stat3 Agonists and Antagonists and Therapeutic Uses Thereof
;  FILE REFERENCE: 10873-009-999
;  CURRENT APPLICATION NUMBER: US/10/380,020
;  CURRENT FILING DATE:  2003-03-07
;  PRIOR APPLICATION NUMBER: 60/231,212
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 769
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-380-020-5

  Query Match             98.6%;  Score 1156;  DB 15;  Length 769;
  Best Local Similarity   98.7%;  Pred. No. 2.8e-93;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 12
US-10-045-792-8
; Sequence 8, Application US/10045792
; Publication No. US20030003563A1
;  GENERAL INFORMATION:
;       APPLICANT: Vinkemeier, Uwe
;                  Moarefi, Ismail
;                  Darnell, Jr., James E.
;                  Kuriyan, John
;       TITLE OF INVENTION: A CRYSTAL OF THE N-TERMINAL DOMAIN OF A
;                           STAT PROTEIN AND METHODS OF USE THEREOF
;       NUMBER OF SEQUENCES: 13
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Klauber & Jackson
;            STREET: 411 Hackensack Avenue, 4th Floor
;            CITY: Hackensack
;            STATE: New Jersey
;            COUNTRY: USA
;            ZIP: 07601
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/045,792
;            FILING DATE: 19-Oct-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/012,710
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Jackson Esq., David A.
;            REGISTRATION NUMBER: 26,742
;            REFERENCE/DOCKET NUMBER: 600-1-194
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 201-487-5800
;            TELEFAX: 201-343-1684
;            TELEX: 133521
;  INFORMATION FOR SEQ ID NO: 8:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 770 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       HYPOTHETICAL: NO
;       SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-045-792-8

  Query Match             98.6%;  Score 1156;  DB 14;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 2.9e-93;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358
```

```
Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 13
US-10-038-010-56
; Sequence 56, Application US/10038010
; Publication No. US20030040089A1
; GENERAL INFORMATION:
;  APPLICANT: HYBRIGENICS
;  APPLICANT:  Pierre, Legrain
;  TITLE OF INVENTION: Protein-protein interactions in adipocyte cells
;  FILE REFERENCE: B4767A
;  CURRENT APPLICATION NUMBER: US/10/038,010
;  CURRENT FILING DATE:  2002-07-23
;  PRIOR APPLICATION NUMBER: US 60/259,377
;  PRIOR FILING DATE: 2001-01-02
;  NUMBER OF SEQ ID NOS: 67
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: STAT3 : Transcription factor
;   LOCATION: (1)..(770)
;   OTHER INFORMATION:
US-10-038-010-56

  Query Match             98.6%;  Score 1156;  DB 14;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 2.9e-93;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 14
US-10-117-087-2
; Sequence 2, Application US/10117087
; Publication No. US20030166854A1
; GENERAL INFORMATION:
;  APPLICANT: SERLUPI-CRESCENZI, Ottaviano
;  APPLICANT:  DELLA PIETRA, Linda
;  TITLE OF INVENTION: ALLELIC VARIANTS OF HUMAN STAT3
;  FILE REFERENCE: SERLUPI=2
;  CURRENT APPLICATION NUMBER: US/10/117,087
;  CURRENT FILING DATE:  2002-04-08
;  PRIOR APPLICATION NUMBER: US/09/526,542
;  PRIOR FILING DATE: 2000-03-19
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Human
US-10-117-087-2

  Query Match             98.6%;  Score 1156;  DB 14;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 2.9e-93;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 15
US-09-925-302-780
; Sequence 780, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA104
;  CURRENT APPLICATION NUMBER: US/09/925,302
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05918
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 896
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 780
;   LENGTH: 793
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-925-302-780

  Query Match             98.6%;  Score 1156;  DB 9;  Length 793;
  Best Local Similarity   98.7%;  Pred. No. 3e-93;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        153 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 212

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        213 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 272

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 332

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 381


Search completed: May 25, 2005, 18:21:54
Job time : 101.62 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:24:07 ; Search time 21.6038 Seconds
(without alignments)
1019.898 Million cell updates/sec

Title: US-10-090-185-30
Perfect score: 1172
Sequence: 1 NHPTAAVVTEKQQMLEQHAQ.....IKTGVQFTTKVRLLVKFPEL 229

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1156	98.6	770	2	I49508	ISGF3 p91-related
2	1151	98.2	770	2	A54444	DNA-binding protei
3	526.5	44.9	739	2	A46159	interferon-depende
4	511	43.6	748	2	A56047	gamma-interferon a
5	321	27.4	851	2	A46160	interferon alpha-i
6	268.5	22.9	786	2	I49274	mammary gland fact
7	268.5	22.9	793	2	S54772	mammary gland fact
8	262.5	22.4	794	2	G02317	transcription acti
9	237	20.2	794	2	S55527	mammary gland fact
10	124.5	10.6	848	2	A54740	interleukin-4-indu
11	121	10.3	978	2	A70387	conserved hypothet
12	119.5	10.2	837	2	I57557	DNA-Binding Protei
13	119.5	10.2	2094	2	S33124	tpr protein - huma
14	119	10.2	533	2	G72593	hypothetical prote
15	118.5	10.1	1166	2	T27075	hypothetical prote
16	118.5	10.1	1208	2	AE1947	chromosome segrega
17	116	9.9	1164	2	T24806	hypothetical prote
18	115	9.8	289	2	S51193	epimorphin - mouse
19	113.5	9.7	1690	2	T13030	microtubule bindin
20	113	9.6	1976	2	A59252	myosin heavy chain
21	113	9.6	1999	1	S21801	myosin heavy chain
22	113	9.6	3187	2	JC5837	364K Golgi complex
23	112.5	9.6	1957	2	T38077	hypothetical coile
24	112	9.6	3685	1	A27605	dystrophin, muscle
25	111.5	9.5	747	1	A57107	kinesin-related pr
26	111.5	9.5	750	2	T38435	coiled coil protei
27	111.5	9.5	857	2	S33821	median body protei
28	111.5	9.5	1178	2	S30431	MSP-300 protein -
29	110.5	9.4	896	2	S43074	epidermal growth f
30	110.5	9.4	924	2	S06117	myosin heavy chain
31	110	9.4	1959	1	A33977	myosin heavy chain
32	109.5	9.3	1992	2	A47297	myosin heavy chain
33	109.5	9.3	2007	1	B43402	myosin heavy chain
34	108.5	9.3	727	2	AC1814	hypothetical prote
35	108.5	9.3	1972	1	A41604	myosin heavy chain
36	108	9.2	464	2	H90279	microtubule bindin
37	108	9.2	734	2	T27055	hypothetical prote
38	108	9.2	1269	2	F84730	probable myosin he
39	107.5	9.2	946	2	S28061	SCP1 protein - rat
40	107.5	9.2	1008	2	AE2304	hypothetical prote
41	107.5	9.2	1509	1	A27224	myosin heavy chain
42	107.5	9.2	1940	2	A59287	myosin heavy chain
43	107.5	9.2	1963	1	MWKW	myosin heavy chain
44	107.5	9.2	2253	2	T30336	nuclear/mitotic ap
45	107	9.1	564	2	A60115	M protein precurso

ALIGNMENTS

RESULT 1
I49508
ISGF3 p91-related transcription factor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I49508; I49009
R;Akira, S.; Nishio, Y.; Inoue, M.; Wang, X.J.; Wei, S.; Matsusaka, T.; Yoshida, K.; Su
Cell 77, 63-71, 1994
A;Title: Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related tr
A;Reference number: A54444; MUID:94208062; PMID:7512451
A;Accession: I49508
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-770 <RES>
A;Cross-references: UNIPROT:P42227; GB:L29278; NID:g476715; PIDN:AAA37254.1; PID:g47671
R;Raz, R.; Durbin, J.E.; Levy, D.E.
J. Biol. Chem. 269, 24391-24395, 1994
A;Title: Acute phase response factor and additional members of the interferon-stimulate
A;Reference number: I49009; MUID:95014185; PMID:7523373
A;Accession: I49009
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-393,'M',395-700,702-770 <RE2>
A;Cross-references: EMBL:U08378; NID:g473889; PIDN:AAA56668.1; PID:g473890
C;Genetics:
A;Gene: APRF
C;Superfamily: human signal transducer and transcription activator STAT5A

Query Match 98.6%; Score 1156; DB 2; Length 770;
Best Local Similarity 98.7%; Pred. No. 1.8e-69;
Matches 226; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358
```

RESULT 2
A54444
DNA-binding protein APRF - human

```
C;Species: Homo sapiens (man)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 09-Jul-2004
C;Accession: A54444
R;Akira, S.; Nishio, Y.; Inoue, M.; Wang, X.J.; Wei, S.; Matsusaka, T.; Yoshida, K.; Sud
Cell 77, 63-71, 1994
A;Title: Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related tra
A;Reference number: A54444; MUID:94208062; PMID:7512451
A;Accession: A54444
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-770 <RES>
A;Cross-references: UNIPROT:P40763; GB:L29277; NID:g475788; PID:g475789
C;Genetics:
A;Gene: GDB:STAT3; APRF
A;Cross-references: GDB:358950
A;Map position: 17q21-17q21
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: DNA binding; transcription factor

  Query Match             98.2%;  Score 1151;  DB 2;  Length 770;
  Best Local Similarity   98.3%;  Pred. No. 4e-69;
  Matches  225;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELHQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 3
A46159
interferon-dependent positive-acting transcription factor ISGF-3 91K chain - human
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A46159
R;Schindler, C.; Fu, X.Y.; Improta, T.; Aebersold, R.; Darnell Jr., J.E.
Proc. Natl. Acad. Sci. U.S.A. 89, 7836-7839, 1992
A;Title: Proteins of transcription factor ISGF-3: one gene encodes the 91-and 84-kDa ISG
A;Reference number: A46159; MUID:92366557; PMID:1502203
A;Accession: A46159
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid; protein
A;Residues: 1-739 <SCH>
A;Cross-references: UNIPROT:P42224
A;Experimental source: HeLa cells
A;Note: sequence extracted from NCBI backbone (NCBIP:110818)
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             44.9%;  Score 526.5;  DB 2;  Length 739;
  Best Local Similarity   46.2%;  Pred. No. 1.3e-27;
  Matches  104;  Conservative   48;  Mismatches   60;  Indels   13;  Gaps    2;

Qy          5 AAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQ 64
              : |: :||: |:   ::|: :|  :| ::| :|:|||::||  |||  |    : ||  :
Db        132 STVMLDKQKELDSKVRNVKDKVMCIEHEIKSLEDLQDEYDFKCKTL--QNREHETNGVAK 189

Qy         65 SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEIACIG 124
              |  :|:   |::|   ||  |: :| ::  ||:  |  |  | ::|| :|||| : ||||
Db        190 SDQKQEQLLLKKMYLMLDNKRKEVVHKIIELLNVTELTQNALINDELVEWKRRQQSACIG 249

Qy        125 GPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERIVELF 184
              |||| |||:|:           | |||:||||||:|| :|: ||| ::: :| :|   ||

Db        250 GPPNACLDQLQ----------QVRQQLKKLEELEQKYTYEHDPITKNKQVLWDRTFSLF 298

Qy        185 RNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              : |::|:|||||||||| || ||||:||||||| |:|||||  ||
Db        299 QQLIQSSFVVERQPCMPTHPQRPLVLKTGVQFTVKLRLLVKLQEL 343


RESULT 4
A56047
gamma-interferon activation site-binding protein Stat4 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004
C;Accession: A56047
R;Yamamoto, K.; Quelle, F.W.; Thierfelder, W.E.; Kreider, B.L.; Gilbert, D.J.; Jenkins,
Mol. Cell. Biol. 14, 4342-4349, 1994
A;Title: Stat4, a novel gamma interferon activation site-binding protein expressed in e
A;Reference number: A56047; MUID:94277038; PMID:8007943
A;Accession: A56047
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-748 <YAM>
A;Cross-references: UNIPROT:P42228; GB:U09351; NID:g509502; PIDN:AAA19692.1; PID:g50950
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: DNA binding; phosphoprotein

  Query Match             43.6%;  Score 511;  DB 2;  Length 748;
  Best Local Similarity   45.5%;  Pred. No. 1.4e-26;
  Matches  102;  Conservative   48;  Mismatches   66;  Indels    8;  Gaps    2;

Qy          8 VTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKS--QGDMQDLNGNNQS 65
              |:|:|: :|     ::  ||  ||  |.:|:|||:||: |||:::  |||       |
Db        136 VSERQRNVEHKVSAIKNSVQMTEQDTKYLEDLQDEFDYRYKTIQTMDQGD------KNSI 189

Qy         66 VTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEIACIGG 125
              :  |::  |::|| :||  |:  :|::  :::  : :  ::  ||| |||:| :||||||
Db        190 LVNQEVLTLQEMLNSLDFKRKEALSKMTQIVNETDLLMNSMLLEELQDWKKRQQIACIGG 249

Qy        126 PPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERIVELFR 185
              | :  ||:|:|  | ||||  | |||::||:|   |::|:||||   |  | ||   |
Db        250 PLHNGLDQLQNCFTLLAESLFQLRQQLEKLQEQSTKMTYEGDPIPAQRAHLLERATFLIY 309

Qy        186 NLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              || |::|||||||||| || ||:|:|| :||| |:|||:| |||
Db        310 NLFKNSFVVERQPCMPTHPQRPMVLKTLIQFTVKLRLLIKLPEL 353


RESULT 5
A46160
interferon alpha-induced transcription activator ISGF-3, 113K chain - human
N;Alternate names: stat2 protein
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A46160; S71908; S53873
R;Fu, X.Y.; Schindler, C.; Improta, T.; Aebersold, R.; Darnell Jr., J.E.
Proc. Natl. Acad. Sci. U.S.A. 89, 7840-7843, 1992
A;Title: The proteins of ISGF-3, the interferon alpha-induced transcriptional activator
A;Reference number: A46160; MUID:92366558; PMID:1502204
A;Accession: A46160
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA; protein
A;Residues: 1-851 <FU1>
A;Cross-references: UNIPROT:P52630
A;Note: sequence extracted from NCBI backbone (NCBIP:110820)
R;Yan, R.; Qureshi, S.; Zhong, Z.; Wen, Z.; Darnell, J.E.
submitted to the EMBL Data Library, December 1994
A;Reference number: S71908
A;Accession: S71908
A;Molecule type: DNA
A;Residues: 1-851 <YAN>
A;Cross-references: EMBL:U18671; NID:g1293919; PIDN:AAA98760.1; PID:g1293920
R;Yan, R.; Qureshi, S.; Zhong, Z.; Wen, Z.; Darnell Jr., J.E.
Nucleic Acids Res. 23, 459-463, 1995
A;Title: The genomic structure of the STAT genes: multiple exons in coincident sites in
```

```
A;Reference number: S53873; MUID:95192056; PMID:7885841
A;Accession: S53873
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-196;392-591;684-730 <YAW>
A;Cross-references: EMBL:U18671
C;Genetics:
A;Gene: stat2
A;Introns: 44/2; 95/3; 127/3; 157/3; 183/1; 211/3; 261/2; 314/2; 345/2; 365/2; 372/2; 40
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: signal transduction; transcription regulation

  Query Match             27.4%;  Score 321;  DB 2;  Length 851;
  Best Local Similarity   33.5%;  Pred. No. 7e-14;
  Matches   74;  Conservative   51;  Mismatches   90;  Indels    6;  Gaps    4;

Qy          8 VTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVT 67
              |  :|  :|    :|  :: | : :  ::: || | | || ::::|    |: :
Db        138 VESQQHEIESRILDLRAMMEKLVKSISQLKDQQDVFCFRYK-IQAKGKTPSLDPH--QTK 194

Qy         68 RQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEIACIGGPP 127
               ||:  |:: |  ||: |: ::     ||  :  : : |   :| :|| : : |||  |
Db        195 EQKI--LQETLNELDKRRKEVLDASKALLGRLTTLIELLL-PKLEEWKAQQQKACIRAPI 251

Qy        128 NICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNL 187
              :  |::|| | |: |:     || :|:|: |   |||: ||: :   :   :: || : |
Db        252 DHGLEQLETWFTAGAKLLFHLRQLLKELKGLSCLVSYQDDPLTKGVDLRNAQVTELLQRL 311

Qy        188 MKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPE 228
              :  ||||| |||||  | |||::||| :|| : ||||:  |
Db        312 LHRAFVVETQPCMPQTPHRPLILKTGSKFTVRTRLLVRLQE 352


RESULT 6
I49274
mammary gland factor - mouse
N;Alternate names: STAT5 protein homolog p80
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I49274; S54773; S54727
R;Liu, X.; Robinson, G.W.; Gouilleux, F.; Groner, B.; Hennighausen, L.
Proc. Natl. Acad. Sci. U.S.A. 92, 8831-8835, 1995
A;Title: Cloning and expression of Stat5 and an additional homologue (Stat5b) involved i
A;Reference number: I49273; MUID:96004632; PMID:7568026
A;Accession: I49274
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-786 <RES>
A;Cross-references: UNIPROT:P42232; UNIPROT:Q9JKM1; EMBL:U21110; NID:g747973; PIDN:AAC52
R;Mui, A.L.F.; Wakao, H.; O'Farrell, A.M.; Harada, N.; Miyajima, A.
EMBO J. 14, 1166-1175, 1995
A;Title: Interleukin-3, granulocyte-macrophage colony stimulating factor and interleukin
A;Reference number: S54772; MUID:95237198; PMID:7720707
A;Accession: S54773
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-432,'E',434-786 <MUI>
A;Cross-references: EMBL:Z48539; NID:g758635; PIDN:CAA88420.1; PID:g758636
R;Azam, M.; Erdjument-Bromage, H.; Kreider, B.L.; Xia, M.; Quelle, F.; Basu, R.; Saris,
EMBO J. 14, 1402-1411, 1995
A;Title: Interleukin-3 signals through multiple isoforms of Stat5.
A;Reference number: S54725; MUID:95246733; PMID:7537213
A;Accession: S54727
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-432,'E',434-786 <AZA>
C;Genetics:
A;Gene: Stat5b
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             22.9%;  Score 268.5;  DB 2;  Length 786;
  Best Local Similarity   30.6%;  Pred. No. 2e-10;
  Matches   71;  Conservative   42;  Mismatches   98;  Indels   21;  Gaps    5;

Qy          5 AAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNY-KTLKSQGDMQDLNGNN 63
              |  :::|   : |  :::|    || | ::| ::  |: |    | ::|: |      |  |
Db        134 ADAMSQKHLQINQTFEELRLITQDTENELKKLQQTQEYFIIQYQESLRIQAQFAQLGQLN 193

Qy         64 -------QSVTRQKMQQLEQML----TALDQMRRSIVSELAGLLSAMEYVQKTLTDEELA 112
                     ::  :||   ||  |       | | |  :  :    |  :   |  : |:||
Db        194 PQERMSRETALQQKQVSLETWLQREAQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELI 253

Qy        113 DWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQH 172
              |||| ::|  ||||   || |::|   |||    | ||||:: | | |::   | |: :
Db        254 QWKRRQQLAGNGGPPEGSLDVLQSWCEKLAEIIWQNRQQIRRAEHLCQQLPIPG-PVEEM 312

Qy        173 RPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLV 224
              :   | ::   |: | |::|:||        | |:||  :|   |||||
Db        313 LAEVNATITDIISALVTSTFIIEKQP--------PQVLKTQTKFAATVRLLV 356


RESULT 7
S54772
mammary gland factor - mouse
N;Alternate names: stat5 protein
C;Species: Mus musculus (house mouse)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S54772; I49273
R;Mui, A.L.F.; Wakao, H.; O'Farrell, A.M.; Harada, N.; Miyajima, A.
EMBO J. 14, 1166-1175, 1995
A;Title: Interleukin-3, granulocyte-macrophage colony stimulating factor and interleuki
A;Reference number: S54772; MUID:95237198; PMID:7720707
A;Accession: S54772
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-793 <MUI>
A;Cross-references: UNIPROT:P42230; UNIPROT:Q9JIA0; EMBL:Z48538; NID:g758633; PIDN:CAA8
R;Liu, X.; Robinson, G.W.; Gouilleux, F.; Groner, B.; Hennighausen, L.
Proc. Natl. Acad. Sci. U.S.A. 92, 8831-8835, 1995
A;Title: Cloning and expression of Stat5 and an additional homologue (Stat5b) involved
A;Reference number: I49273; MUID:96004632; PMID:7568026
A;Accession: I49273
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-793 <RES>
A;Cross-references: EMBL:U21103; NID:g747971; PIDN:AAA80590.1; PID:g747972
C;Genetics:
A;Gene: Stat5a
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             22.9%;  Score 268.5;  DB 2;  Length 793;
  Best Local Similarity   30.6%;  Pred. No. 2e-10;
  Matches   70;  Conservative   42;  Mismatches   96;  Indels   21;  Gaps    5;

Qy          8 VTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNY-KTLKSQGDMQDLNGNN--- 63
              :::|   : |  :::|    || | ::| ::  |: |    | ::|: |      |   |
Db        137 MSQKHLQINQRFEELRLITQDTENELKKLQQTQEYFIIQYQESLRIQAQFAQLGQLNPQE 196

Qy         64 ----QSVTRQKMQQLEQML----TALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWK 115
                 ::  :||   ||  |       | | |  :  :    |  :   |  : |:||  ||
Db        197 RMSRETALQQKQVSLETWLQREAQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELIQWK 256

Qy        116 RRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPM 175
              || ::|  ||||   || |::|   |||   | ||||:: | | |::   | |: :
Db        257 RRQQLAGNGGPPEGSLDVLQSWCEKLAEIIWQNRQQIRRAEHLCQQLPIPG-PVEEMLAE 315

Qy        176 LEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLV 224
              :   | ::   |: | |::|:||        | |:||  :|   |||||
Db        316 VNATITDIISALVTSTFIIEKQP--------PQVLKTQTKFAATVRLLV 356


RESULT 8
G02317
transcription activator stat5A - human
C;Species: Homo sapiens (man)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 09-Jul-2004
```

```
C;Accession: G02317
R;Lin, J.
submitted to the EMBL Data Library, December 1995
A;Reference number: H01043
A;Accession: G02317
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-794 <LIN>
A;Cross-references: UNIPROT:P42229; EMBL:U43185; NID:g1151169; PIDN:AAB06589.1; PID:g115
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             22.4%;  Score 262.5;  DB 2;  Length 794;
  Best Local Similarity   30.1%;  Pred. No. 5.1e-10;
  Matches   69;  Conservative   43;  Mismatches   96;  Indels   21;  Gaps    5;

Qy          8 VTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNY-KTLKSQGDMQDL------- 59
              :::|   : |  :::|   || | ::| ::  |: |   | ::|: |     |
Db        137 MSQKHLQINQTFEELRLVTQDTENELKKLQQTQEYFIIQYQESLRIQAQFAQLAQLSPQE 196

Qy         60 NGNNQSVTRQKMQQLEQML----TALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWK 115
                : ::  :||   ||  |       | | |  :  :     |  :   |  : |:||  ||
Db        197 RLSRETALQQKQVSLEAWLQREAQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELIQWK 256

Qy        116 RRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPM 175
              || ::|  ||||   || |::|   |||   | ||||:: | | |::   | |: :
Db        257 RRQQLAGNGGPPEGSLDVLQSWCEKLAEIIWQNRQQIRRAEHLCQQLPIPG-PVEEMLAE 315

Qy        176 LEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLV 224
              :   | ::   |: | |::|:||        | |:||  :|   |||||
Db        316 VNATITDIISALVTSTFIIEKQP--------PQVLKTQTKFAATVRLLV 356


RESULT 9
S55527
mammary gland factor - sheep
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S55527; S44353
R;Wakao, H.; Gouilleux, F.; Groner, B.
EMBO J. 14, 854-855, 1995
A;Title: Corrigenda. Mammary gland factor (MGF) is a novel member of the cytokine regula
A;Reference number: S55527; MUID:95188889; PMID:7882987
A;Accession: S55527
A;Molecule type: mRNA
A;Residues: 1-794 <WAK>
A;Cross-references: UNIPROT:P42231; EMBL:X78428; NID:g602354; PIDN:CAA55191.1; PID:g6023
A;Note: this is a revision to the sequence from reference S44353
R;Wakao, H.; Gouilleux, F.; Groner, B.
EMBO J. 13, 2182-2191, 1994
A;Title: Mammary gland factor (MGF) is a novel member of the cytokine regulated transcri
A;Reference number: S44353; MUID:94244619; PMID:7514531
A;Accession: S44353
A;Molecule type: mRNA
A;Residues: 1-716,'RHLHGPGSLPSR',729,'P',731,'ASL' <WAW>
A;Cross-references: EMBL:X78428
A;Note: this sequence has been revised in reference S55527
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             20.2%;  Score 237;  DB 2;  Length 794;
  Best Local Similarity   29.3%;  Pred. No. 2.5e-08;
  Matches   67;  Conservative   42;  Mismatches   98;  Indels   22;  Gaps    6;

Qy          8 VTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNY-KTLKSQGDMQDLNGNN--- 63
              :::|   : |  :::|   || | ::| ::  |: |   | ::|: |   .  |   |
Db        138 MSQKHLQINQTFEELRLVTQDTENELKKLQQTQEYFIIQYQESLRIQAQFAQLAQLNPQE 197

Qy         64 ----QSVTRQKMQQLEQML----TALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWK 115
                  ::  :||   ||  |       | | |  :  :     |  :   |  : |:||  ||
Db        198 RLSRETALQQKQVSLEAWLQREAQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELIQWK 257

Qy        116 RRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPM 175
              || :   :  ||   || |::|   |||   | ||||:: | | |::   | |: :
Db        258 RRHDWRGMEAPPR-SLDVLQSWCEKLAEIIWQNRQQIRRAEHLCQQLPIPG-PVEEMLAE 315

Qy        176 LEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLV 224
              :   | ::   |: | |::|:||        | |:||  :|   |||||
Db        316 VNATITDIISALVTSTFIIEKQP--------PQVLKTQTKFAATVRLLV 356


RESULT 10
A54740
interleukin-4-induced transcription factor stat - human
C;Species: Homo sapiens (man)
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004
C;Accession: A54740
R;Hou, J.; Schindler, U.; Henzel, W.J.; Ho, T.C.; Brasseur, M.; McKnight, S.L.
Science 265, 1701-1706, 1994
A;Title: An interleukin-4-induced transcription factor: IL-4 stat.
A;Reference number: A54740; MUID:94367369; PMID:8085155
A;Accession: A54740
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-848 <HOU>
A;Cross-references: UNIPROT:P42226
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: DNA binding; transcription regulation

  Query Match             10.6%;  Score 124.5;  DB 2;  Length 848;
  Best Local Similarity   29.7%;  Pred. No. 0.82;
  Matches   41;  Conservative   22;  Mismatches   52;  Indels   23;  Gaps    4;

Qy         88 IVSELAGLLSAMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQ 147
              :: |  | | |      | |  : :  |||: ::|  | |           :  |||  | :
Db        182 LLQETTGELEA----AKALVLKRIQIWKRQQQLAGNGAP----------FEESLAPLQER 227

Qy        148 TRQQIKKLEELQQKVSYKGDPI-VQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDR 206
                    :   :|||:|   |  :  : |  |  |: |: | |: | | |:||:||
Db        228 CESLVDIYSQLQQEVGAAGGELEPKTRASLTGRLDEVLRTLVTSCFLVEKQP-------- 279

Qy        207 PLVIKTGVQFTTKVRLLV 224
              | |:||  :|   || |:
Db        280 PQVLKTQTKFQAGVRFLL 297


RESULT 11
A70387
conserved hypothetical protein aq_1006 - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C;Accession: A70387
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: A70387
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-978 <AQF>
A;Cross-references: UNIPROT:O67124; GB:AE000718; NID:g2983504; PIDN:AAC07092.1; PID:g29
A;Experimental source: strain VF5
C;Genetics:
A;Gene: aq_1006
C;Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

  Query Match             10.3%;  Score 121;  DB 2;  Length 978;
  Best Local Similarity   23.3%;  Pred. No. 1.7;
  Matches   49;  Conservative   38;  Mismatches   81;  Indels   42;  Gaps    6;

Qy         10 EKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQ 69
              |::  ||:    |  ::::||   | || |::  :|: |                :
Db        234 EEKDSLERELSQVVTKLKELENLEKEVEKLREKLEFSRKVAP----------YVPIAK 281

Qy         70 KMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLT---------------DEELAD- 113
              ::::::: || |   :  :  ||| |   : : |: |                 ::||
Db        282 RIEEIDKKLTELKVRKNKLTKELAVLKDELSFAQEELNRIEAEKEKFKEEKEREKELEHR 341
```

```
Qy        114 WKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHR 173
Db        342 LKKLQEIKEI-------LKELSQLSSSLKEKEREYEQAKQEFEDLSERVE-KGKKLVAET 393

Qy        174 PMLEERIVELFRN------LMKSAFVVERQ 197
Db        394 EEKLEKIKELFSEEEYTSLKMKERLLVELQ 423


RESULT 12
I57557
DNA-Binding Protein and transcription factor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C;Accession: I57557
R;Quelle, F.W.; Shimoda, K.; Thierfelder, W.; Fischer, C.L.; Kim, A.; Ruben, S.M.; Cleve
Mol. Cell. Biol. 15, 3336-3343, 1995
A;Title: Cloning of murine Stat6 and human Stat6, Stat proteins that are tyrosine phosph
A;Reference number: I57557; MUID:95280934; PMID:7760829
A;Accession: I57557
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-837 <RES>
A;Cross-references: UNIPROT:P52633; GB:L47650; NID:g1008876; PIDN:AAA79006.1; PID:g10088
C;Genetics:
A;Gene: STAT6
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             10.2%;  Score 119.5;  DB 2;  Length 837;
  Best Local Similarity   25.8%;  Pred. No. 1.7;
  Matches   60;  Conservative   31;  Mismatches   83;  Indels   59;  Gaps   10;

Qy         10 EKQQMLEQ--HAQDVRKRVQD----------LEQKMKVVENLQDDFDFNYKTLKSQGDMQ 57
Db        106 EKKAVIEEFRHLPGPFHRKQEELKFTTPLGRLHHRVRETRLLRESLHLGPKT--GQVSLQ 163

Qy         58 D-----LNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELA 112
Db        164 NLIDPPLNGPGPS---------EDLPTILQ----------GTVGDLETTQ-PLVLLRIQ 202

Qy        113 DWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLE-ELQQKVSYKGDPIVQ 171
Db        203 IWKRQQQLAGNGTPFEESLAGLQERCESLVEIYSQLHQEIGAASGELEPKT--------- 253

Qy        172 HRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLV 224
Db        254 -RASLISRLDEVLRTLVTSSFLVEKQP--------PQVLKTQTKFQAGVRFLL 297


RESULT 13
S33124
tpr protein - human
N;Alternate names: kinase-related transforming protein (tpr-met); protein with promoter
C;Species: Homo sapiens (man)
C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C;Accession: S33124; S23740; S00928; G01185
R;Mitchell, P.J.; Cooper, C.S.
Oncogene 7, 2329-2333, 1992
A;Title: The human tpr gene encodes a protein of 2094 amino acids that has extensive coi
A;Reference number: S33124; MUID:93064711; PMID:1437155
A;Accession: S33124
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-2094 <MIT>
A;Cross-references: UNIPROT:Q15624; UNIPROT:Q9UE33; EMBL:X66397; NID:g633225
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991
R;Mitchell, P.J.; Cooper, C.S.
Oncogene 7, 383-388, 1992
A;Title: Nucleotide sequence analysis of human tpr cDNA clones.
A;Reference number: S23740; MUID:92195670; PMID:1549355
A;Accession: S23740
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-725,'L' <MI2>
A;Cross-references: EMBL:X63105; NID:g37257; PIDN:CAA44819.1; PID:g37258
R;King, H.W.S.; Tempest, P.R.; Merrifield, K.R.; Rance, A.J.
Oncogene 2, 617-619, 1988
A;Title: Tpr homologues activate met and raf.
A;Reference number: S00928; MUID:88262257; PMID:3387099
A;Accession: S00928
A;Molecule type: mRNA
A;Residues: 1-31,'R',33-142 <KIN>
A;Cross-references: EMBL:Y00672; NID:g37255; PIDN:CAA68681.1; PID:g37256
R;Greco, A.
submitted to the EMBL Data Library, December 1995
A;Reference number: H00592
A;Accession: G01185
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 144-228 <GRE>
A;Cross-references: EMBL:X94208; NID:g1296797; PIDN:CAA63904.1; PID:g1296798
C;Genetics:
A;Gene: GDB:TPR
A;Cross-references: GDB:128821; OMIM:189940
A;Map position: 1q25-1q25
A;Introns: 177/3

  Query Match             10.2%;  Score 119.5;  DB 2;  Length 2094;
  Best Local Similarity   20.4%;  Pred. No. 5.1;
  Matches   45;  Conservative   50;  Mismatches   81;  Indels   45;  Gaps    6;

Qy          6 AVVTEKQQMLEQHAQDVRK-------------RVQDLEQKMKV---VENLQDDFDFNYK 48
Db       1387 ASLTNNQNLIQSLKEDLNKVRTEKETIQKDLDAKIIDIQEKVKTITQVKKIGRRYKTQYE 1446

Qy         49 TLKSQ-----------GDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLL 96
Db       1447 ELKAQQDKVMETSAQSSGDHQE-----QHVSVQEMQELKETLNQAETKSKSLESQ----- 1496

Qy         97 SAMEYVQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLE 156
Db       1497 --VENLQKTLSEKETEARNLQEQTVQLQSE----LSRLRQDLQDRTTQEEQLRQQITEKE 1550

Qy        157 ELQQKVSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQ 197
Db       1551 EKTRKAIVAAKSKIAHLAGVKDQLTKENEELKQRNGALDQQ 1591


RESULT 14
G72593
hypothetical protein APE1216 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: G72593
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takai
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: G72593
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-533 <KAW>
A;Cross-references: UNIPROT:Q9YCP2; DDBJ:AP000061; NID:g5104821; PIDN:BAA80205.1; PID:d
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1216

  Query Match             10.2%;  Score 119;  DB 2;  Length 533;
  Best Local Similarity   20.7%;  Pred. No. 1.1;
  Matches   41;  Conservative   49;  Mismatches   66;  Indels   42;  Gaps    7;

Qy          8 VTEKQQMLEQHAQDVRKRVQDLEQKMKVVEN----LQDDF---------------DFNYK 48
Db        319 MSQQLQALAEDLESLSSRVEDLEARVGSVEDRLSQAEEDIDSLTTSLDSLRTELEDLSTR 378
```

```
Qy          49 TLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTD 108
               ::|  ::|||     |    :|||:| |   ::  :::  :|| | : :| :|::: :
Db         379 LAEAQASLEDLNTRLDQVA-STLQQLQQRLATAEESLQALTEDLASLQAEVETLQQSIVE 437

Qy         109 EELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQ----TRQQIKKLEELQQKVSY 164
                :    : |  :  :         |||  : || |  :|     ::|   :|: | ::
Db         438 IDRRLGQLRSTVDAV---------RLE--VESLGEKLVQAEEKNQRQDASIEDFQSQIQ- 485

Qy         165 KGDPIVQHRPMLEERIVE 182
                     : |  |:|:  |
Db         486 ------ELRTQLDEKTRE 497


RESULT 15
T27075
hypothetical protein Y51A2D.16 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T27075
R;McMurray, A.
submitted to the EMBL Data Library, January 1998
A;Reference number: Z20307
A;Accession: T27075
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1166 <WIL>
A;Cross-references: EMBL:AL021497; PIDN:CAA16403.1; GSPDB:GN00023; CESP:Y51A2D.16
A;Experimental source: clone Y51A2D
C;Genetics:
A;Gene: CESP:Y51A2D.16
A;Map position: 5
A;Introns: 17/3; 45/1; 76/2; 156/3; 240/2; 288/2; 594/2; 657/3; 756/3; 798/2; 900/3; 948

  Query Match             10.1%;  Score 118.5;  DB 2;  Length 1166;
  Best Local Similarity   20.7%;  Pred. No. 3;
  Matches   50;  Conservative   47;  Mismatches   75;  Indels   69;  Gaps    8;

Qy           1 NHPTAAVVTEKQ----QMLEQHAQ-----------DVRKRVQDLEQKMKVVENLQDDFD 44
               :| | :: :||     ::||   |            :  ||::| :| :  : ||:.:: :
Db         430 SHVTRSLSSEKNTLKAKLLELQDQVEAQTLELLNQKNCGKRLEDRDQMISNLHNLKNELE 489

Qy          45 FNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEY--- 101
               : || ::| :::           :|:|:| : |       |  ::  :  | : :
Db         490 NDLKTCQTQLELES----------KKLQRLREDLVLEKSRRADLIGRIHSLCTTLSLNGA 539

Qy         102 -VQKTLTDEELAD-----------WKRRPEIACIGGPPNI------------------ 129
                :|   |:|| |              ||  :   | |   |
Db         540 NFEKINNDDELIDNIDDIMMNALVAVKRERDDLRIQGNQQIQELHDLKRDIEKLRRSESE 599

Qy         130 CLDRLENWITSLAESQLQTRQQI----KKLEELQQKVSYKGDPIVQHRPMLEERIVELFR 185
                |:  :: :  |    : |::|:    :|| ||  ::| | | |      |::  | || |
Db         600 SLNESDDRVRELTRENMHTKEQVFMLQEKLRELNLELSTKNDEI----DMVKASIEELNR 655

Qy         186 N 186
               |
Db         656 N 656


Search completed: May 25, 2005, 17:45:29
Job time : 22.6038 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:15:30 ; Search time 95.284 Seconds
(without alignments)
1230.701 Million cell updates/sec

Title: US-10-090-185-30
Perfect score: 1172
Sequence: 1 NHPTAAVVTEKQQMLEQHAQ.....IKTGVQFTTKVRLLVKFPEL 229

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1156	98.6	770	1	STA3_HUMAN	P40763 homo sapien
2	1156	98.6	770	1	STA3_MOUSE	P42227 mus musculu
3	1153	98.4	770	1	STA3_RAT	P52631 rattus norv
4	1149	98.0	770	1	STA3_BOVIN	P61635 bos taurus
5	1131	96.5	771	2	Q6DV79	Q6dv79 gallus gall
6	1098	93.7	769	2	Q9PVX8	Q9pvx8 xenopus lae
7	1095	93.4	766	2	Q7ZXK3	Q7zxk3 xenopus lae
8	1010	86.2	414	2	Q7ZTS5	Q7zts5 brachydanio
9	1010	86.2	765	2	Q6DVF3	Q6dvf3 oryzias lat
10	1010	86.2	785	2	Q6GUE7	Q6gue7 oryzias lat
11	1010	86.2	786	2	Q6NV46	Q6nv46 brachydanio
12	1010	86.2	806	2	O93599	O93599 brachydanio
13	986	84.1	764	2	Q90Y16	Q90y16 tetraodon f
14	976	83.3	767	2	O13133	O13133 oncorhynchu
15	656	56.0	163	2	Q9N145	Q9n145 macaca mula
16	578	49.3	751	2	Q8JGN0	Q8jgn0 xenopus lae
17	576	49.1	712	2	Q6P6Q7	Q6p6q7 rattus norv
18	576	49.1	749	2	Q9QXK0	Q9qxk0 rattus norv
19	574	49.0	749	2	Q8C497	Q8c497 mus musculu
20	573	48.9	712	2	Q99K94	Q99k94 mus musculu
21	573	48.9	749	2	Q8C3V4	Q8c3v4 mus musculu
22	573	48.9	749	2	Q9D323	Q9d323 mus musculu
23	573	48.9	755	2	Q8C8M3	Q8c8m3 mus musculu
24	572	48.8	750	1	STA1_HUMAN	P42224 homo sapien
25	572	48.8	750	2	Q68D00	Q68d00 homo sapien
26	572	48.8	757	2	Q764M5	Q764m5 sus scrofa
27	565	48.2	749	1	STA1_MOUSE	P42225 mus musculu
28	541.5	46.2	754	2	O13131	O13131 oncorhynchu
29	535	45.6	749	2	O93598	O93598 brachydanio
30	535	45.6	749	2	Q6P943	Q6p943 brachydanio
31	534	45.6	528	2	Q8JFU8	Q8jfu8 brachydanio
32	532.5	45.4	108	2	Q704W6	Q704w6 bos taurus
33	530.5	45.3	718	2	Q801Y2	Q801y2 carassius a
34	523	44.6	748	1	STA4_HUMAN	Q14765 homo sapien
35	520.5	44.4	758	2	Q90Y17	Q90y17 tetraodon f
36	518.5	44.2	754	2	O13132	O13132 oncorhynchu
37	516	44.0	657	2	Q8AW24	Q8aw24 brachydanio
38	507.5	43.3	749	1	STA4_MOUSE	P42228 mus musculu
39	507	43.3	651	2	Q7ZZ53	Q7zz53 brachydanio
40	507	43.3	667	2	Q8AW20	Q8aw20 brachydanio
41	506.5	43.2	652	2	Q7ZZ77	Q7zz77 brachydanio
42	506	43.2	553	2	Q8JFU7	Q8jfu7 brachydanio
43	506	43.2	748	2	Q66HB2	Q66hb2 rattus norv
44	504	43.0	1153	2	Q8JFS5	Q8jfs5 brachydanio
45	494	42.2	737	2	Q90Y15	Q90y15 tetraodon f

ALIGNMENTS

RESULT 1
STA3_HUMAN
ID STA3_HUMAN STANDARD; PRT; 770 AA.
AC P40763; O14916; Q9BW54;
DT 01-FEB-1995 (Rel. 31, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Signal transducer and activator of transcription 3 (Acute-phase
DE response factor).
GN Name=STAT3; Synonyms=APRF;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Placenta;
RX MEDLINE=94208062; PubMed=7512451; DOI=10.1016/0092-8674(94)90235-6;
RA Akira S., Nishio Y., Inoue M., Wang X.-J., Wei S., Matsusaka T.,
RA Yoshida K., Sudo T., Naruto M., Kishimoto T.;
RT "Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-
RT related transcription factor involved in the gp130-mediated signaling
RT pathway.";
RL Cell 77:63-71(1994).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=98296260; PubMed=9630560; DOI=10.1016/S0378-1119(98)00185-1;
RA Della Pietra L., Bressan A., Pezzotti A., Serlupi-Crescenzi O.;
RT "Highly conserved amino-acid sequence between murine STAT3 and a
RT revised human STAT3 sequence.";
RL Gene 213:119-124(1998).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM DEL-701), AND VARIANT ILE-143.
RA Rieder M.J., Daniels R.L., da Ponte S.H., Hastings N.C., Ahearn M.O.,
RA Rajkumar N., Yi Q., Nickerson D.A.;
RT "SeattleSNPs. NHLBI HL66682 program for genomic applications, UW-
RT FHCRC, Seattle, WA (URL: http://pga.gs.washington.edu).";
RL Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND DEL-701).
RC TISSUE=Kidney, and Pancreas;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,

```
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   SEQUENCE OF 564-704 FROM N.A.
RC   TISSUE=Liver;
RA   Della Pietra L., Bressan A., Pezzotti A.R., Serlupi-Crescenzi O.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   PHOSPHORYLATION ON SERINE.
RX   MEDLINE=95215843; PubMed=7701321;
RA   Zhang X., Blenis J., Li H.-C., Schindler C., Chen-Kiang S.;
RT   "Requirement of serine phosphorylation for formation of STAT-promoter
RT   complexes.";
RL   Science 267:1990-1994(1995).
RN   [7]
RP   INTERACTION WITH NCOA1.
RX   PubMed=11773079; DOI=10.1074/jbc.M111486200;
RA   Giraud S., Bienvenu F., Avril S., Gascan H., Heery D.M., Coqueret O.;
RT   "Functional interaction of STAT3 transcription factor with the
RT   coactivator NcoA/SRC1a.";
RL   J. Biol. Chem. 277:8004-8011(2002).
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes.
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P40763-1; Sequence=Displayed;
CC       Name=Del-701;
CC         IsoId=P40763-2; Sequence=VSP_010474;
CC   -!- TISSUE SPECIFICITY: Heart, brain, placenta, lung, liver, skeletal
CC       muscle, kidney and pancreas.
CC   -!- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,
CC       LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation
CC       is important for the formation of stable DNA-binding STAT3
CC       homodimers and maximal transcriptional activity.
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L29277; AAA58374.1; -.
DR   EMBL; AJ012463; CAA10032.1; -.
DR   EMBL; AY572796; AAS66986.1; -.
DR   EMBL; BC000627; AAH00627.1; -.
DR   EMBL; BC014482; AAH14482.1; -.
DR   EMBL; AF029311; AAB84254.1; -.
DR   PIR; A54444; A54444.
DR   HSSP; P42227; 1BG1.
DR   TRANSFAC; T01493; -.
DR   Genew; HGNC:11364; STAT3.
DR   H-InvDB; HIX0013840; -.
DR   MIM; 102582; -.
DR   GO; GO:0005737; C:cytoplasm; TAS.
DR   GO; GO:0005634; C:nucleus; TAS.
DR   GO; GO:0005062; F:hematopoietin/interferon-class (D200-domain. . .; TAS.
DR   GO; GO:0003700; F:transcription factor activity; TAS.
DR   GO; GO:0006928; P:cell motility; TAS.
DR   GO; GO:0007259; P:JAK-STAT cascade; TAS.
DR   GO; GO:0000122; P:negative regulation of transcription from P. . .; TAS.
DR   GO; GO:0007399; P:neurogenesis; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
KW   Activator; Alternative splicing; DNA-binding; Nuclear protein;
KW   Phosphorylation; Polymorphism; SH2 domain; Transcription regulation.
FT   DOMAIN      580    670       SH2.
FT   MOD_RES     705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES     727    727       Phosphoserine (By similarity).
FT   VARSPLIC    701    701       Missing (in isoform Del-701).
FT                                /FTId=VSP_010474.
FT   VARIANT      32     32       Q -> K (in dbSNP:1803125).
FT                                /FTId=VAR_018683.
FT   VARIANT     143    143       M -> I.
FT                                /FTId=VAR_018679.
FT   CONFLICT    288    288       Q -> H (in Ref. 1).
FT   CONFLICT    460    460       P -> S (in Ref. 1).
FT   CONFLICT    548    548       K -> N (in Ref. 1).
FT   CONFLICT    561    561       F -> Y (in Ref. 1).
FT   CONFLICT    667    667       V -> L (in Ref. 1).
FT   CONFLICT    730    730       T -> A (in Ref. 1).
SQ   SEQUENCE   770 AA;  88067 MW;  6C00632211C8012D CRC64;

  Query Match             98.6%;  Score 1156;  DB 1;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 3.1e-65;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 2
STA3_MOUSE
ID   STA3_MOUSE     STANDARD;      PRT;   770 AA.
AC   P42227;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Signal transducer and activator of transcription 3 (Acute-phase
DE   response factor).
GN   Name=Stat3; Synonyms=Aprf;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A), AND SEQUENCE OF 154-158; 181-185
RP   AND 632-640.
RC   STRAIN=BALB/c; TISSUE=Liver;
RX   MEDLINE=94208062; PubMed=7512451; DOI=10.1016/0092-8674(94)90235-6;
```

```
RA   Akira S., Nishio Y., Inoue M., Wang X.J., Shi W., Matsusaka T.,
RA   Yoshida K., Sudo T., Naruto M., Kishimoto T.;
RT   "Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-
RT   related transcription factor involved in the gp130-mediated signaling
RT   pathway.";
RL   Cell 77:63-71(1994).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   TISSUE=Thymus;
RX   MEDLINE=94188718; PubMed=8140422;
RA   Zhong Z., Wen Z., Darnell J.E. Jr.;
RT   "Stat3: a STAT family member activated by tyrosine phosphorylation in
RT   response to epidermal growth factor and interleukin-6.";
RL   Science 264:95-98(1994).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   TISSUE=Brain;
RX   MEDLINE=95014185; PubMed=7523373;
RA   Raz R., Durbin J.E., Levy D.E.;
RT   "Acute phase response factor and additional members of the interferon-
RT   stimulated gene factor 3 family integrate diverse signals from
RT   cytokines, interferons, and growth factors.";
RL   J. Biol. Chem. 269:24391-24395(1994).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3B).
RC   STRAIN=BALB/c, and C57BL/6; TISSUE=Liver;
RX   MEDLINE=96016116; PubMed=7568080;
RA   Schaefer T.S., Sanders L.K., Nathans D.;
RT   "Cooperative transcriptional activity of Jun and Stat3 beta, a short
RT   form of Stat3.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:9097-9101(1995).
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   STRAIN=129/SvJ;
RX   PubMed=11161808; DOI=10.1006/geno.2000.6433;
RA   Miyoshi K., Cui Y., Riedlinger G., Lehoczky J., Zon L., Oka T.,
RA   Dewar K., Hennighausen L.;
RT   "Structure of the mouse stat 3/5 locus: evolution from Drosophila to
RT   zebrafish to mouse.";
RL   Genomics 71:150-155(2001).
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   STRAIN=C57BL/6J, and NOD/LtJ;
RA   Davoodi-Semiromi A., She J.-X.;
RT   "A mutant Stat5b with weaker DNA binding defines a key defective
RT   pathway in non-obese diabetic (NOD) mice.";
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   STRAIN=FVB/N; TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [8]
RP   PHOSPHORYLATION SITE SER-727, AND MUTAGENESIS.
RX   MEDLINE=95354205; PubMed=7543024; DOI=10.1016/0092-8674(95)90311-9;
RA   Wen Z., Zhong Z., Darnell J.E. Jr.;
RT   "Maximal activation of transcription by Stat1 and Stat3 requires both
RT   tyrosine and serine phosphorylation.";
RL   Cell 82:241-250(1995).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 136-716.
RX   MEDLINE=98334373; PubMed=9671298; DOI=10.1038/28101;
RA   Becker S., Groner B., Mueller C.W.;
RT   "Three-dimensional structure of the Stat3beta homodimer bound to
RT   DNA.";
RL   Nature 394:145-151(1998).
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes. STAT3B interacts with the N-terminal
CC       part of JUN to activate such promoters in a cooperative way.
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=Stat3A;
CC         IsoId=P42227-1; Sequence=Displayed;
CC       Name=Stat3B;
CC         IsoId=P42227-2; Sequence=VSP_006287;
CC       Name=Del-701;
CC         IsoId=P42227-3; Sequence=VSP_010475;
CC   -!- TISSUE SPECIFICITY: STAT3A is seen in the liver, spleen, and
CC       kidney. STAT3B is also detected in the liver, although in a much
CC       less abundant manner.
CC   -!- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,
CC       LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation
CC       is important for the formation of stable DNA-binding STAT3
CC       homodimers and maximal transcriptional activity (By similarity).
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L29278; AAA37254.1; -.
DR   EMBL; U06922; AAA19452.1; -.
DR   EMBL; U08378; AAA56668.1; -.
DR   EMBL; U30709; AAC52612.1; -.
DR   EMBL; AF246978; AAL59017.1; -.
DR   EMBL; AY299489; AAQ75418.1; -.
DR   EMBL; AY299490; AAQ75419.1; -.
DR   EMBL; BC003806; AAH03806.1; -.
DR   PIR; I49508; I49508.
DR   PDB; 1BG1; X-ray; A=1-722.
DR   TRANSFAC; T01574; -.
DR   MGD; MGI:103038; Stat3.
DR   GO; GO:0005737; C:cytoplasm; IDA.
DR   GO; GO:0005634; C:nucleus; IDA.
DR   GO; GO:0005886; C:plasma membrane; IDA.
DR   GO; GO:0003677; F:DNA binding; IDA.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   GO; GO:0016563; F:transcriptional activator activity; IDA.
DR   GO; GO:0007259; P:JAK-STAT cascade; IDA.
DR   GO; GO:0006357; P:regulation of transcription from Pol II pro. . .; IDA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
```

```
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
KW   3D-structure; Activator; Acute phase; Alternative splicing;
KW   Direct protein sequencing; DNA-binding; Nuclear protein;
KW   Phosphorylation; SH2 domain; Transcription regulation.
FT   DOMAIN      580    670       SH2.
FT   MOD_RES     705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES     727    727       Phosphoserine.
FT   VARSPLIC    716    770       TTCSNTIDLPMSPRTLDSLMQFGNNGEGAEPSAGGQFESLT
FT                                FDMDLTSECATSPM -> FIDAVWK (in isoform
FT                                Stat3B).
FT                                /FTId=VSP_006287.
FT   VARSPLIC    701    701       Missing (in isoform Del-701).
FT                                /FTId=VSP_010475.
FT   MUTAGEN     727    727       S->A: Decreased transcriptional
FT                                activation.
FT   CONFLICT     16     16       E -> K (in Ref. 2).
FT   CONFLICT     25     25       S -> T (in Ref. 2 and 4).
FT   CONFLICT    394    394       M -> I (in Ref. 1).
FT   HELIX       139    180
FT   TURN        181    182
FT   TURN        197    198
FT   HELIX       199    237
FT   TURN        238    238
FT   HELIX       239    251
FT   TURN        252    253
FT   HELIX       261    290
FT   TURN        294    295
FT   TURN        297    301
FT   HELIX       302    320
FT   STRAND      321    328
FT   TURN        330    331
FT   TURN        333    334
FT   TURN        336    337
FT   STRAND      338    340
FT   TURN        341    342
FT   STRAND      345    351

  Query Match             98.6%;  Score 1156;  DB 1;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 3.1e-65;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 3
STA3_RAT
ID   STA3_RAT       STANDARD;      PRT;   770 AA.
AC   P52631;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Signal transducer and activator of transcription 3.
GN   Name=Stat3;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=96102059; PubMed=8530402; DOI=10.1074/jbc.270.50.29998;
RA   Ripperger J.A., Fritz S., Richter K., Hocke G.M., Lottspeich F.,
RA   Fey G.H.;
RT   "Transcription factors Stat3 and Stat5b are present in rat liver
RT   nuclei late in an acute phase response and bind interleukin-6 response
RT   elements.";
RL   J. Biol. Chem. 270:29998-30006(1995).
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes.
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation (By similarity).
CC   -!- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,
CC       LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation
CC       is important for the formation of stable DNA-binding STAT3
CC       homodimers and maximal transcriptional activity (By similarity).
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X91810; CAA62920.1; -.
DR   HSSP; P42227; 1BG1.
DR   RGD; 3772; Stat3.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
KW   Activator; DNA-binding; Nuclear protein; Phosphorylation; SH2 domain;
KW   Transcription regulation.
FT   DOMAIN      580    670       SH2.
FT   MOD_RES     705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES     727    727       Phosphoserine (By similarity).
SQ   SEQUENCE   770 AA;  88039 MW;  D74A0C76954754ED CRC64;

  Query Match             98.4%;  Score 1153;  DB 1;  Length 770;
  Best Local Similarity   98.3%;  Pred. No. 4.8e-65;
  Matches  225;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |:||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 4
```

```
STA3_BOVIN
ID   STA3_BOVIN     STANDARD;      PRT;   770 AA.
AC   P61635;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Signal transducer and activator of transcription 3.
GN   Name=STAT3;
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Mammary gland;
RA   Seyfert H.M., Wheeler T.T., Moolenaar A., Pitra C.;
RT   "The STAT5B-encoding gene was flipped across the STAT3/STAT5A-locus
RT   during ruminant evolution.";
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes (By similarity).
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation (By similarity).
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ620655; CAF06182.1; -.
DR   PROSITE; PS50001; SH2; 1.
KW   Activator; DNA-binding; Nuclear protein; Phosphorylation; SH2 domain;
KW   Transcription regulation.
FT   DOMAIN       580    670       SH2.
FT   MOD_RES      705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES      727    727       Phosphoserine (By similarity).
SQ   SEQUENCE   770 AA;  87974 MW;  9CEB147C73E83274 CRC64;

  Query Match             98.0%;  Score 1149;  DB 1;  Length 770;
  Best Local Similarity   98.3%;  Pred. No. 8.6e-65;
  Matches  225;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||| |||||||||||||||| ||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLLDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 5
Q6DV79
ID   Q6DV79      PRELIMINARY;      PRT;   771 AA.
AC   Q6DV79;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Signal transducer and activator of transcription 3.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zhou G.Y., Leung F.C.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY641397; AAT64887.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   771 AA;  88174 MW;  71AC855C5DEC03E2 CRC64;

  Query Match             96.5%;  Score 1131;  DB 2;  Length 771;
  Best Local Similarity   96.9%;  Pred. No. 1.2e-63;
  Matches  221;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 HPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNG 61
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        131 HPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNG 190

Qy         62 NNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEIA 121
              |||||||||||||||||||||||||| |||||||||||||| | |||||||||| :||
Db        191 NNQSVTRQKMQQLEQMLTALDQMRRGIVSELAGLLSAMEYVQKMLADEELADWKRRQQIA 250

Qy        122 CIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERIV 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 CIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERIV 310

Qy        182 ELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||| ||||||||||||||||
Db        311 ELFRNLMKSAFVVERQPCMPMHPDRPLVFKTGVQFTTKVRLLVKFPEL 358


RESULT 6
Q9PVX8
ID   Q9PVX8      PRELIMINARY;      PRT;   769 AA.
AC   Q9PVX8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat 3.
GN   Name=stat 3;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20107399; PubMed=10642787; DOI=10.1006/dbio.1999.9518;
RA   Nishinakamura R., Matsumoto Y., Matsuda T., Ariizumi T., Heike T.,
RA   Asashima M., Yokota T.;
RT   "Activation of Stat3 by cytokine receptor gp130 ventralizes Xenopus
RT   embryos independent of BMP-4.";
RL   Dev. Biol. 216:481-490(1999).
DR   EMBL; AB017701; BAA86061.1; -.
```

```
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   769 AA;  87974 MW;  0905C03263303069 CRC64;

  Query Match             93.7%;  Score 1098;  DB 2;  Length 769;
  Best Local Similarity   92.6%;  Pred. No. 1.5e-61;
  Matches  212;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              :|| |||||||||||||| |||||:|||||||||||||||||||||||||| |: :||
Db        130 SHPNAAVVTEKQQMLEQHLQDVRKKVQDLEQKMKVVENLQDDFDFNYKTLKSQSDLSELN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||:||:|:|:|| |||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQLRRTIISDLASLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQIRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||| ||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTNKVRLLVKFPEL 358


RESULT 7
Q7ZXK3
ID   Q7ZXK3      PRELIMINARY;      PRT;   766 AA.
AC   Q7ZXK3;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat3-A protein.
GN   Name=stat3-A;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC044717; AAH44717.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   766 AA;  87599 MW;  31018A3321CCEB9C CRC64;

  Query Match             93.4%;  Score 1095;  DB 2;  Length 766;
  Best Local Similarity   91.7%;  Pred. No. 2.3e-61;
  Matches  210;  Conservative   11;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              :|| |||:|||||||||| |||||:|||||||||||||||||||||||| |: :||
Db        130 SHPNAAVITEKQQMLEQHLQDVRKKVQDLEQKMKVVENLQDDFDFNYKTLKSQSDLSELN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||:||:|:|:|| |||||||||| ||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQLRRTIISDLASLLSAMEYVQKNLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              |||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIRKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||| ||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTNKVRLLVKFPEL 358


RESULT 8
Q7ZTS5
ID   Q7ZTS5      PRELIMINARY;      PRT;   414 AA.
AC   Q7ZTS5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat3 protein.
GN   Name=stat3;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB; TISSUE=Whole body;
```

```
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB; TISSUE=Whole body;
RA   Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC045276; AAH45276.1; -.
DR   HSSP; P42227; 1BG1.
DR   ZFIN; ZDB-GENE-980526-68; stat3.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
SQ   SEQUENCE   414 AA;  48253 MW;  0FFD1B509B7526BD CRC64;

  Query Match             86.2%;  Score 1010;  DB 2;  Length 414;
  Best Local Similarity   84.3%;  Pred. No. 2.9e-56;
  Matches  194;  Conservative   21;  Mismatches   13;  Indels    2;  Gaps    2;

Qy          2 HPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLN 60
              |||  |||||||:|| : ||:||||||:|||||::|||||||||||||||||:: ||||
Db        130 HPTGTVVTEKQQILEHNLQDIRKRVQDMEQKMKMLENLQDDFDFNYKTLKSAGELSQDLN 189

Qy         61 GNNQ-SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPE 119
              ||:| : ||||| |||||:||||:|| ||:|:|||||||::||| |||||||||||| :
Db        190 GNSQAAATRQKMSQLEQMLSALDQLRRQIVTEMAGLLSAMDFVQKNLTDEELADWKRRQQ 249

Qy        120 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 179
              |||||||||||||| |||||||||| ||||:||||||||||||:|||| |||:
Db        250 IACIGGPPNICLDRLETWITSLAESQLQIRQQIRKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy        180 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              ||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        310 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 359


RESULT 9
Q6DVF3
ID   Q6DVF3      PRELIMINARY;      PRT;   765 AA.
AC   Q6DVF3;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Signal transducer and activation of transcription factor 3.
OS   Oryzias latipes (Medaka fish) (Japanese ricefish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Liu R., Hong Y.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY639947; AAT64912.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   765 AA;  87566 MW;  F5D01408748EC703 CRC64;

  Query Match             86.2%;  Score 1010;  DB 2;  Length 765;
  Best Local Similarity   84.3%;  Pred. No. 5.6e-56;
  Matches  194;  Conservative   20;  Mismatches   14;  Indels    2;  Gaps    2;

Qy          2 HPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLN 60
              |||  |||||||:|| : ||:||||||:|||||::|||||||||||||||||:: ||||
Db        130 HPTGTVVTEKQQILEHNLQDIRKRVQDMEQKMKMLENLQDDFDFNYKTLKSQGELNQDLN 189

Qy         61 GNNQ-SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPE 119
              ||:| : ||||| |||||:||||:|| ||:|: |||:||:|||| |||||||||||| :
Db        190 GNSQAAATRQKMAQLEQMLSALDQLRRQIVTEMGGLLTAMDYVQKNLTDEELADWKRRQQ 249

Qy        120 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 179
              |||||||||||||| |||||||||| ||||||||||||||||||||:|||| |||:
Db        250 IACIGGPPNICLDRLETWITSLAESQLQIRQQIKKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy        180 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              ||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        310 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTNKVRLLVKFPEL 359


RESULT 10
Q6GUE7
ID   Q6GUE7      PRELIMINARY;      PRT;   785 AA.
AC   Q6GUE7;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Signal transducer and activator of transcription 3 isoform 1.
OS   Oryzias latipes (Medaka fish) (Japanese ricefish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Liu R., Hong Y.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY641434; AAT46364.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
```

```
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   785 AA;  89643 MW;  81F231BDE27DE938 CRC64;

  Query Match             86.2%;  Score 1010;  DB 2;  Length 785;
  Best Local Similarity   84.3%;  Pred. No. 5.7e-56;
  Matches  194;  Conservative   20;  Mismatches   14;  Indels    2;  Gaps    2;

Qy          2 HPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLN 60
              |||  |||||||:|| : ||:||||||:|||||::||||||||||||||:: ||||
Db        130 HPTGTVVTEKQQILEHNLQDIRKRVQDMEQKMKMLENLQDDFDFNYKTLKSQGELNQDLN 189

Qy         61 GNNQ-SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPE 119
              ||:| : ||||| |||||:||||:|| ||:|: |||:||:|||| |||||||||||| :
Db        190 GNSQAAATRQKMAQLEQMLSALDQLRRQIVTEMGGLLTAMDYVQKNLTDEELADWKRRQQ 249

Qy        120 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 179
              |||||||||||||| |||||||||| ||||||||||||||||||:|||| |||:
Db        250 IACIGGPPNICLDRLETWITSLAESQLQIRQQIKKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy        180 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              ||:||||||||||||||||||||||||||||||||||| ||||||||||
Db        310 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTNKVRLLVKFPEL 359


RESULT 11
Q6NV46
ID   Q6NV46      PRELIMINARY;      PRT;   786 AA.
AC   Q6NV46;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Stat3 protein.
GN   Name=stat3;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC068320; AAH68320.1; -.
DR   ZFIN; ZDB-GENE-980526-68; stat3.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   786 AA;  90039 MW;  FC7371D0B0E5447E CRC64;

  Query Match             86.2%;  Score 1010;  DB 2;  Length 786;
  Best Local Similarity   84.3%;  Pred. No. 5.7e-56;
  Matches  194;  Conservative   21;  Mismatches   13;  Indels    2;  Gaps    2;

Qy          2 HPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLN 60
              |||  |||||||:|| : ||:||||||:|||||::||||||||||||||:: |:: ||||
Db        130 HPTGTVVTEKQQILEHNLQDIRKRVQDMEQKMKMLENLQDDFDFNYKTLKSAGELSQDLN 189

Qy         61 GNNQ-SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPE 119
              ||:| : ||||| |||||:||||:|| ||:|:|||||||::||| |||||||||||| :
Db        190 GNSQAAATRQKMSQLEQMLSALDQLRRQIVTEMAGLLSAMDFVQKNLTDEELADWKRRQQ 249

Qy        120 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 179
              |||||||||||||| |||||||||| ||||:||||||||||||:|||| |||:
Db        250 IACIGGPPNICLDRLETWITSLAESQLQIRQQIRKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy        180 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              ||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        310 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 359


RESULT 12
O93599
ID   O93599      PRELIMINARY;      PRT;   806 AA.
AC   O93599;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Transcription factor.
GN   Name=stat3;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Oates A.C.;
RL   Thesis (1998), University of Melbourne, Australia.
DR   EMBL; AJ005693; CAA06677.1; -.
DR   HSSP; P42227; 1BG1.
DR   ZFIN; ZDB-GENE-980526-68; stat3.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   806 AA;  92151 MW;  74BC4EA401C3C942 CRC64;

  Query Match             86.2%;  Score 1010;  DB 2;  Length 806;
  Best Local Similarity   84.3%;  Pred. No. 5.9e-56;
```

```
  Matches  194;  Conservative   21;  Mismatches   13;  Indels    2;  Gaps    2;

Qy          2 HPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLN 60
              |||  |||||||:|| : ||:||||||:|||||::|||||||||||||||| |:: ||||
Db        130 HPTGTVVTEKQQILEHNLQDIRKRVQDMEQKMKMLENLQDDFDFNYKTLKSAGELSQDLN 189

Qy         61 GNNQ-SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPE 119
              ||:| : ||||| |||||:||||:|| ||:|:|||||||::||| ||||||||||| :
Db        190 GNSQAAATRQKMSQLEQMLSALDQLRRQIVTEMAGLLSAMDFVQKNLTDEELADWKRRQQ 249

Qy        120 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 179
              |||||||||||||||| |||||||||| ||||:||||||||||||||:|||| |||:
Db        250 IACIGGPPNICLDRLETWITSLAESQLQIRQQIRKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy        180 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              ||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        310 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 359


RESULT 13
Q90Y16
ID   Q90Y16      PRELIMINARY;      PRT;   764 AA.
AC   Q90Y16;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   STAT3.
OS   Tetraodon fluviatilis (Puffer fish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=47145;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sung S.C., Fan T.J., Leu J.H., Chou C.M., Huang C.J.;
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF307106; AAL09415.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   764 AA;  87411 MW;  E661FFE18BEFD8BE CRC64;

  Query Match             84.1%;  Score 986;  DB 2;  Length 764;
  Best Local Similarity   83.0%;  Pred. No. 1.8e-54;
  Matches  191;  Conservative   21;  Mismatches   16;  Indels    2;  Gaps    2;

Qy          2 HPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLN 60
              :|:  |||||||||| : ||:||||||:|||||::|||||||||||||||:: ||||
Db        130 NPSGTVVTEKQQMLELNLQDIRKRVQDMEQKMKMLENLQDDFDFNYKTLKSQGELSQDLN 189

Qy         61 GNNQ-SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPE 119
              ||:| |  |||| ||||||:||||:|| ||:|: |||:||:|||| ||:||||||| :
Db        190 GNSQASAARQKMAQLEQMLSALDQLRRQIVTEMGGLLTAMDYVQKNLTNEELADWKRRQQ 249

Qy        120 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 179
              |||||||||||||||| |||||||||| ||||||||||||||||||:|||| |||:
Db        250 IACIGGPPNICLDRLETWITSLAESQLQIRQQIKKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy        180 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              ||:|||||||||||||||||||||||||||||||||| |||||||| ||
Db        310 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTNKVRLLVKFQEL 359


RESULT 14
O13133
ID   O13133      PRELIMINARY;      PRT;   767 AA.
AC   O13133;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat3.
GN   Name=rbtStat3;
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8022;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Johnson M.C., Mourich D.V., Leong J.C.;
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U60333; AAB60926.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
SQ   SEQUENCE   767 AA;  87816 MW;  FED97740C74C3798 CRC64;

  Query Match             83.3%;  Score 976;  DB 2;  Length 767;
  Best Local Similarity   81.0%;  Pred. No. 8e-54;
  Matches  187;  Conservative   25;  Mismatches   17;  Indels    2;  Gaps    2;

Qy          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDL 59
              :||:  |||||||:|| : ||:||||||:|||||::|||||||||:||||||:: ||:
Db        129 SHPSGTVVTEKQQILEHNLQDIRKRVQDMEQKMKMLENLQDDFDFHYKTLKSQGELSQDM 188

Qy         60 NGNNQ-SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRP 118
              |||:| : ||||| ||||||:||||:|| ||:|:|||||||::||| |||:||||||||
Db        189 NGNSQAAATRQKMSQLEQMLSALDQLRRQIVTEMAGLLSAMDFVQKNLTDDELADWKRRQ 248

Qy        119 EIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEE 178
              :|||||||| ||||||| ||||| | ||| |||||||||||||||||||:|||| |||
Db        249 QIACIGGPPKICLDRLETWITSLGEIQLQIRQQIKKLEELQQKVSYKGDPIIQHRPALEE 308

Qy        179 RIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              :||:|||||||||||||| |||||||||||||||| ||||||||||
Db        309 KIVDLFRNLMKSAFVVERQRCMPMHPDRPLVIKTGVQFTNKVRLLVKFPEL 359


RESULT 15
Q9N145
ID   Q9N145      PRELIMINARY;      PRT;   163 AA.
AC   Q9N145;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Signal transducer and activator of transcription (Fragment).
GN   Name=STAT3;
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Arredondo J.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; AF227560; AAF73401.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
FT   NON_TER       1      1
FT   NON_TER     163    163
SQ   SEQUENCE   163 AA;  18855 MW;  DD3F11E7F5B9878C CRC64;

  Query Match             56.0%;  Score 656;  DB 2;  Length 163;
  Best Local Similarity   98.4%;  Pred. No. 2.9e-34;
  Matches  126;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        102 VQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQK 161
              ||||||||||||||| :||||||||||||||||||||||||||||||||||||||||||||
Db          1 VQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQK 60

Qy        162 VSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVR 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVR 120

Qy        222 LLVKFPEL 229
              ||||||||
Db        121 LLVKFPEL 128


Search completed: May 25, 2005, 17:43:44
Job time : 96.284 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:14:45 ; Search time 121.891 Seconds
(without alignments)
726.619 Million cell updates/sec

Title: US-10-090-185-31
Perfect score: 1172
Sequence: 1 NHPTAAVVTEKQQMLEQHLQ.....IKTGVQFTTKVRLLVKFPEL 229

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1172	100.0	229	4	AAY72863	Aay72863 Mouse Sta
2	1168	99.7	229	4	AAY72850	Aay72850 Mouse Sta
3	1168	99.7	252	4	AAY72846	Aay72846 Mouse Sta
4	1168	99.7	271	4	AAY72841	Aay72841 Mouse Sta
5	1168	99.7	770	2	AAR72082	Aar72082 Mouse Sta
6	1168	99.7	770	2	AAW03176	Aaw03176 Mouse STA
7	1163	99.2	229	4	AAY72862	Aay72862 Mouse Sta
8	1157.5	98.8	228	4	AAY72861	Aay72861 Mouse Sta
9	1157	98.7	720	5	AAE22055	Aae22055 Human Sta
10	1157	98.7	769	5	ABB57164	Abb57164 Mouse isc
11	1157	98.7	769	5	AAE22054	Aae22054 Human Sta
12	1157	98.7	769	5	AAE22056	Aae22056 Human pro
13	1157	98.7	770	2	AAR82995	Aar82995 Mouse liv
14	1157	98.7	770	2	AAY03768	Aay03768 Human STA
15	1157	98.7	770	3	AAB12377	Aab12377 N-termina
16	1157	98.7	770	5	AAE14652	Aae14652 Murine ST
17	1157	98.7	770	5	ABG69497	Abg69497 Human bai
18	1157	98.7	770	6	ABU10476	Abu10476 Mouse STA
19	1157	98.7	770	8	ADN04365	Adn04365 Antipsori
20	1157	98.7	770	8	ADP54789	Adp54789 Human PRO
21	1157	98.7	793	3	AAB58442	Aab58442 Lung canc
22	1154	98.5	770	7	ADD44738	Add44738 Rat Prote
23	1152	98.3	770	2	AAR82993	Aar82993 Human pla
24	1152	98.3	770	4	AAB19964	Aab19964 Human sig
25	1152	98.3	770	5	AAE15174	Aae15174 Human Sta
26	1152	98.3	770	7	ADD44740	Add44740 Human Pro
27	1150	98.1	229	4	AAY72860	Aay72860 Mouse Sta
28	1090	93.0	213	4	AAY72851	Aay72851 Mouse Sta
29	1090	93.0	236	4	AAY72847	Aay72847 Mouse Sta
30	1044	89.1	223	4	AAY72854	Aay72854 Mouse Sta
31	845	72.1	185	4	AAY72855	Aay72855 Mouse Sta
32	777	66.3	176	4	AAY72848	Aay72848 Mouse Sta
33	653	55.7	128	4	AAY72852	Aay72852 Mouse Sta
34	601	51.3	143	4	AAY72849	Aay72849 Mouse Sta
35	573	48.9	749	5	AAG78526	Aag78526 Rat STAT-
36	571	48.7	129	4	AAY72856	Aay72856 Mouse Sta
37	569	48.5	268	4	AAY72844	Aay72844 Mouse Sta
38	569	48.5	582	2	AAW62996	Aaw62996 Human tru
39	569	48.5	582	6	ABU04748	Abu04748 Human exp
40	569	48.5	582	8	ADH57036	Adh57036 Truncated
41	569	48.5	680	6	ABR59713	Abr59713 Human sig
42	569	48.5	712	2	AAR72079	Aar72079 Human Sta
43	569	48.5	712	2	AAW03170	Aaw03170 Human STA
44	569	48.5	712	2	AAW62995	Aaw62995 Human Sta
45	569	48.5	712	6	ABU04747	Abu04747 Human exp

ALIGNMENTS

RESULT 1
AAY72863
ID AAY72863 standard; protein; 229 AA.
XX
AC AAY72863;
XX
DT 31-MAY-2001 (first entry)
XX
DE Mouse Stat3 mutant (V151A) protein fragment (130-358 amino acids).
XX
KW Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW cellular transformation; dysproliferative disease; cancer; psoriasis;
KW therapy; mutant; mutein.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT Region 1. .25
FT /note= "Stat3-c-Jun interaction region 1; corresponds to
FT 130-154 position of Stat3 protein"
FT Misc-difference 22
FT /note= "Wild type Val substituted with Ala corresponds to
FT 151 position of Stat-3 protein"
FT Region 213. .229
FT /note= "Stat3-c-Jun interaction region 2; corresponds to
FT 342-358 position of Stat3 protein"
XX
PN WO200116605-A2.
XX
PD 08-MAR-2001.
XX
PF 30-AUG-2000; 2000WO-US023822.
XX
PR 31-AUG-1999; 99US-00387418.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI Zhang X, Horvath C, Wrzeszcynska MH, Darnell JE;
XX
DR WPI; 2001-226705/23.
XX
PT Identifying an agent for use in modulating the interaction between
PT transcription factor c-Jun and a Stat3 protein.
XX
PS Claim 66; Page 86; 86pp; English.
XX
CC The present sequence is mouse Stat3 mutant (V151A) protein fragment
CC containing 130-358 amino acids of Stat3 protein. This mutant is obtained

```
CC   by replacing Val 151 with Ala in Stat3 protein. The invention relates to
CC   methods for identifying interacting regions of transcription factors and
CC   methods for identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 229 AA;

  Query Match           100.0%;  Score 1172;  DB 4;  Length 229;
  Best Local Similarity  100.0%;  Pred. No. 6.5e-100;
  Matches  229;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 2
AAY72850
ID   AAY72850 standard; protein; 229 AA.
XX
AC   AAY72850;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 protein fragment #8 (130-358 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .25
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
FT   Region          213. .229
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 76-77; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 130-358
CC   amino acids of Stat3 protein. This Stat3 fragment showed strong binding
CC   to c-Jun protein in the cell extract. The invention relates to methods
CC   for identifying interacting regions of transcription factors and methods
CC   for identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 229 AA;

  Query Match           99.7%;  Score 1168;  DB 4;  Length 229;
  Best Local Similarity  99.6%;  Pred. No. 1.5e-99;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 3
AAY72846
ID   AAY72846 standard; protein; 252 AA.
XX
AC   AAY72846;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 protein fragment #4 (107-358 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          24. .48
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
FT   Region          236. .252
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
```

```
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 73; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 107-358
CC   amino acids of Stat3 protein. This Stat3 fragment showed strong binding
CC   to c-Jun protein in the cell extract. The invention relates to methods
CC   for identifying interacting regions of transcription factors and methods
CC   for identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 252 AA;

  Query Match             99.7%;  Score 1168;  DB 4;  Length 252;
  Best Local Similarity   99.6%;  Pred. No. 1.7e-99;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db         24 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 83

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 203

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 252


RESULT 4
AAY72841
ID   AAY72841 standard; protein; 271 AA.
XX
AC   AAY72841;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 protein fragment #2 (107-377 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          24. .48
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
FT   Region          236. .252
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 65; Page 67-68; 86pp; English.
XX
CC   The present sequence is mouse Stat3 protein fragment containing 107-377
CC   amino acids of Stat3 protein. This Stat3 fragment showed strong binding
CC   to c-Jun protein in the cell extract. The invention relates to methods
CC   for identifying interacting regions of transcription factors and methods
CC   for identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 271 AA;

  Query Match             99.7%;  Score 1168;  DB 4;  Length 271;
  Best Local Similarity   99.6%;  Pred. No. 1.9e-99;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db         24 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 83

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 203

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 252


RESULT 5
AAR72082
ID   AAR72082 standard; protein; 770 AA.
XX
AC   AAR72082;
XX
DT   25-MAR-2003  (revised)
DT   27-SEP-1995  (first entry)
XX
DE   Mouse Stat3 (19sf6).
XX
KW   Signal transducer and activator of transcription; STAT; 19sf6; Stat3;
KW   receptor recognition factor; transcription factor; cellular debilitation;
KW   derangement; dysfunction; interferon-gamma.
XX
OS   Mus sp.
XX
PN   WO9508629-A1.
XX
PD   30-MAR-1995.
XX
PF   26-SEP-1994;   94WO-US010849.
XX
PR   24-SEP-1993;   93US-00126588.
PR   24-SEP-1993;   93US-00126595.
```

```
PR   11-MAR-1994;   94US-00212184.
PR   11-MAR-1994;   94US-00212185.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Darnell JE,  Schindler CW,  Shuai K,  Wen Z,  Zhong Z;
XX
DR   WPI; 1995-139598/18.
DR   N-PSDB; AAQ89340.
XX
PT   Receptor recognition factor implicated in transcriptional stimulation of
PT   genes - useful in drug screening assays and/or for treating cellular
PT   debilitations, derangements and/or dysfunctions, etc.
XX
PS   Claim 1; Page 107-110; 160pp; English.
XX
CC   A fragment encoding the human Stat91 protein was used to screen a murine
CC   thymus and spleen cDNA for homologous proteins. A highly homologous gene
CC   (given in AAQ89338) was isolated that encoded a 91 kDa protein (AAR72080)
CC   (Stat1) that was responsive to interferon- gamma. Using a fragment of the
CC   mouse gene as probe, 2 additional members of the 113-91 family of
CC   receptor recognition factor proteins were isolated. The 2 genes (AAQ89339
CC   -40) were cloned in plasmids 13sf1 and 19sf6 and encoded proteins termed
CC   Stat4 (AAR72081) and Stat3 (AAR72082), respectively. (Updated on 25-MAR-
CC   2003 to correct PN field.)
XX
SQ   Sequence 770 AA;

  Query Match             99.7%;  Score 1168;  DB 2;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 7.5e-99;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 6
AAW03176
ID   AAW03176 standard; protein; 770 AA.
XX
AC   AAW03176;
XX
DT   24-OCT-1996  (first entry)
XX
DE   Mouse STAT4.
XX
KW   STAT; STAT4; signal transducer and activator of transcription;
KW   DNA binding protein; ligand; receptor; oncogenesis; inflammation;
KW   autoimmune disease; antagonist; therapy.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Domain          398. .508
FT                   /label= DNA_binding_domain
FT                   /note= "Claim 3, page 110"
XX
PN   WO9620954-A2.
XX
PD   11-JUL-1996.
XX
PF   28-DEC-1995;   95WO-US017025.
XX
PR   06-JAN-1995;   95US-00369796.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Darnell JE,  Wen Z,  Horvath CM,  Zhong Z;
XX
DR   WPI; 1996-333941/33.
DR   N-PSDB; AAT31280.
XX
PT   New STAT protein DNA-binding domain peptide(s) - useful for diagnosing,
PT   preventing or treating cellular dysfunction, e.g. oncogenesis,
PT   inflammation, parasitic disease or autoimmunity.
XX
PS   Disclosure; Page 87-90; 138pp; English.
XX
CC   Mouse signal transducer and activator of transcription (STAT) protein
CC   STAT4 (AAW03176) serves a dual purpose, i.e. signal transduction from
CC   ligand-activated receptor kinase complexes followed by nuclear
CC   translocation and DNA binding to activate transcription. Recombinant
CC   STAT4 can be obtd. using cDNA clone 19sf6 (AAT31278) obtd. from
CC   splenic/thymic cells. STAT4 includes a DNA-binding domain (see also
CC   AAW03167) capable of both receptor recognition and message delivery via
CC   DNA binding in a receptor-ligand specific manner. STAT proteins and their
CC   DNA binding domains (see also AAW03165-75) are useful for screening
CC   antagonists used to inhibit STAT-mediated signal transduction and
CC   activation of transcription
XX
SQ   Sequence 770 AA;

  Query Match             99.7%;  Score 1168;  DB 2;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 7.5e-99;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 7
AAY72862
ID   AAY72862 standard; protein; 229 AA.
XX
AC   AAY72862;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 mutant (L148A) protein fragment (130-358 amino acids).
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy; mutant; mutein.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .25
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
```

```
FT   Misc-difference 19
FT                   /note= "Wild type Leu substituted with Ala; corresponds
FT                   to 148 position of Stat-3 protein"
FT   Region          213. .229
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 66; Page 85; 86pp; English.
XX
CC   The present sequence is mouse Stat3 mutant (L148A) protein fragment
CC   containing 130-358 amino acids of Stat3 protein. This mutant is obtained
CC   by replacing Leu 148 with Ala in Stat3 protein. The invention relates to
CC   methods for identifying interacting regions of transcription factors and
CC   methods for identifying agents which modulates the interaction between a
CC   transcription factor such as c-Jun and a Stat protein such as Stat-1 and
CC   Stat-3, useful for modulating gene transcription e.g., cellular
CC   transformation. These identifying agents are used in the treatment of
CC   dysproliferative diseases and also for treating cancer and psoriasis. A
CC   Stat protein comprises the N-terminal domain, coiled-coil domain, DNA
CC   binding domain, linker domain, SH2 domain and transactivation domain
XX
SQ   Sequence 229 AA;

  Query Match             99.2%;  Score 1163;  DB 4;  Length 229;
  Best Local Similarity   99.1%;  Pred. No. 4.4e-99;
  Matches  227;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||| || ||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 8
AAY72861
ID   AAY72861 standard; protein; 228 AA.
XX
AC   AAY72861;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse Stat3 mutant (T346A, K348A, R350A) protein fragment.
XX
KW   Mouse; Stat3 protein; transcription factor; c-Jun; gene transcription;
KW   cellular transformation; dysproliferative disease; cancer; psoriasis;
KW   therapy; mutant; mutein.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          1. .25
FT                   /note= "Stat3-c-Jun interaction region 1; corresponds to
FT                   130-154 position of Stat3 protein"
FT   Region          213. .229
FT                   /note= "Stat3-c-Jun interaction region 2; corresponds to
FT                   342-358 position of Stat3 protein"
FT   Misc-difference 217
FT                   /note= "Wild type Thr substituted with Ala; corresponds
FT                   to 346 position of Stat-3 protein"
FT   Misc-difference 219
FT                   /note= "Wild type Lys substituted with Ala; corresponds
FT                   to 348 position of Stat-3 protein"
FT   Misc-difference 221
FT                   /note= "Wild type Arg substituted with Ala; corresponds
FT                   to 350 position of Stat-3 protein"
XX
PN   WO200116605-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US023822.
XX
PR   31-AUG-1999;   99US-00387418.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Zhang X,  Horvath C,  Wrzeszcynska MH,  Darnell JE;
XX
DR   WPI; 2001-226705/23.
XX
PT   Identifying an agent for use in modulating the interaction between
PT   transcription factor c-Jun and a Stat3 protein.
XX
PS   Claim 66; Page 84-85; 86pp; English.
XX
CC   The present sequence is mouse Stat3 mutant protein fragment containing
CC   130-358 amino acids of Stat3 protein. This mutant is obtained by
CC   replacing Thr 346 with Ala, Lys 348 with Ala and Arg 350 with Ala in the
CC   Stat3 protein. The invention relates to methods for identifying
CC   interacting regions of transcription factors and methods for identifying
CC   agents which modulates the interaction between a transcription factor
CC   such as c-Jun and a Stat protein such as Stat-1 and Stat-3, useful for
CC   modulating gene transcription e.g., cellular transformation. These
CC   identifying agents are used in the treatment of dysproliferative diseases
CC   and also for treating cancer and psoriasis. A Stat protein comprises the
CC   N-terminal domain, coiled-coil domain, DNA binding domain, linker domain,
CC   SH2 domain and transactivation domain
XX
SQ   Sequence 228 AA;

  Query Match             98.8%;  Score 1157.5;  DB 4;  Length 228;
  Best Local Similarity   99.6%;  Pred. No. 1.4e-98;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQD-RKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 59

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         60 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 119

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 179

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        180 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 228

RESULT 9
AAE22055
ID   AAE22055 standard; protein; 720 AA.
XX
AC   AAE22055;
XX
DT   25-JUL-2002  (first entry)
XX
DE   Human Stat3beta protein.
XX
KW   Human; signal transducer and activator of transcription 3; ischaemia;
KW   immune response; Stat3; coronary atherosclerosis; vascular occlusion;
KW   hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;
KW   inflammation; chronic obstructive pulmonary disease; cardiac arrest;
KW   insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;
KW   shock; chronic active hepatitis; adult respiratory distress syndrome;
KW   nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;
KW   Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;
KW   polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;
KW   proliferative disease; Grave's disease; ulcerative colitis; sarcoma;
KW   carcinoma; degenerative disorder; gene therapy; growth deficiency;
KW   cirrhosis; hypoproliferative disorder; lesion; Statbeta.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 713. .714
FT                   /note= "Encoded by ACA CCA TTC"
XX
PN   WO200220032-A1.
XX
PD   14-MAR-2002.
XX
PF   10-SEP-2001; 2001WO-US028254.
XX
PR   08-SEP-2000; 2000US-0231212P.
XX
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (UYSF-) UNIV SOUTH FLORIDA.
XX
PI   Yu H,  Pardoll D,  Jove R,  Dalton W;
XX
DR   WPI; 2002-362218/39.
DR   N-PSDB; AAD35066.
XX
PT   Modulating angiogenesis and an immune response in an individual, for
PT   treating a hypoxic or ischemic condition, comprises administering a
PT   compound that modulates the activity of a signal transducer and activator
PT   of transcription 3.
XX
PS   Disclosure; Page 87-89; 94pp; English.
XX
CC   The invention relates to a method of modulating angiogenesis and immune
CC   response. Method involves administering to an individual a compound that
CC   modulate the activity of signal transducer and activator of transcription
CC   3 (Stat3). Modulating angiogenesis is useful for treating or preventing
CC   hypoxic or ischaemic condition or disorder which is the result of stroke,
CC   ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,
CC   tissue ischaemia in the lower extremities, infarction, trauma, vascular
CC   occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,
CC   chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,
CC   epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest,
CC   nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy
CC   with neovascularisation. Suppressing an immune response is useful for
CC   ameliorating a symptom of an autoimmune disease such as systemic lupus
CC   erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,
CC   Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,
CC   mixed connective tissue disease, primary biliary cirrhosis, pernicious
CC   anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,
CC   gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,
CC   idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's
CC   disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune
CC   infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and
CC   dense deposit disease. The method is useful in preventing or treating
CC   specific proliferative and oncogenic disease which includes sarcomas and
CC   carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,
CC   fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,
CC   hypoproliferative disorders, physical trauma, lesions and wounds. The
CC   method is also used in gene therapy. The present sequence is human
CC   Stat3beta protein
XX
SQ   Sequence 720 AA;

  Query Match             98.7%;  Score 1157;  DB 5;  Length 720;
  Best Local Similarity   98.7%;  Pred. No. 7.2e-98;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 10
ABB57164
ID   ABB57164 standard; protein; 769 AA.
XX
AC   ABB57164;
XX
DT   07-MAR-2002  (first entry)
XX
DE   Mouse ischaemic condition related protein sequence SEQ ID NO:398.
XX
KW   Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW   vasospastic ischaemia; ischaemic condition; ischaemic disease.
XX
OS   Mus musculus.
XX
PN   WO200188188-A2.
XX
PD   22-NOV-2001.
XX
PF   18-MAY-2001; 2001WO-JP004192.
XX
PR   18-MAY-2000; 2000JP-00145977.
XX
PA   (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI   Ishikawa K,  Asai S,  Takahashi Y,  Nagata T,  Ishii Y;
XX
DR   WPI; 2002-034733/04.
DR   N-PSDB; ABI99454.
XX
PT   Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT   expression levels of particular genes defined in the specification or by
PT   determining the expression profile of a gene group comprising these
PT   genes.
XX
PS   Claim 2; Page 1084-1087; 2690pp; English.
XX
CC   The present invention describes a method for examining ischaemic
CC   conditions, comprising measuring the expression levels of particular
```

CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (ABI99202 to ABI99912, encoding the
CC protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The expression
CC levels or expression profiles produced by these genes are used as an
CC indicator when screening for ischaemic condition-improving drugs or
CC therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR
CC primers for a mouse ischaemic condition related sequence, which are used
CC in the exemplification of the present invention
XX
SQ Sequence 769 AA;

Query Match 98.7%; Score 1157; DB 5; Length 769;
Best Local Similarity 98.7%; Pred. No. 7.8e-98;
Matches 226; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358
```

RESULT 11
AAE22054
ID AAE22054 standard; protein; 769 AA.
XX
AC AAE22054;
XX
DT 25-JUL-2002 (first entry)
XX
DE Human Stat3 protein.
XX
KW Human; signal transducer and activator of transcription 3; ischaemia;
KW immune response; Stat3; coronary atherosclerosis; vascular occlusion;
KW hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;
KW inflammation; chronic obstructive pulmonary disease; cardiac arrest;
KW insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;
KW shock; chronic active hepatitis; adult respiratory distress syndrome;
KW nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;
KW Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;
KW polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;
KW proliferative disease; Grave's disease; ulcerative colitis; sarcoma;
KW carcinoma; degenerative disorder; gene therapy; growth deficiency;
KW cirrhosis; hypoproliferative disorder; lesion.
XX
OS Homo sapiens.
XX
PN WO200220032-A1.
XX
PD 14-MAR-2002.
XX
PF 10-SEP-2001; 2001WO-US028254.
XX
PR 08-SEP-2000; 2000US-0231212P.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
PA (UYSF-) UNIV SOUTH FLORIDA.
XX
PI Yu H, Pardoll D, Jove R, Dalton W;
XX
DR WPI; 2002-362218/39.
DR N-PSDB; AAD35065.
XX
PT Modulating angiogenesis and an immune response in an individual, for
PT treating a hypoxic or ischemic condition, comprises administering a
PT compound that modulates the activity of a signal transducer and activator
PT of transcription 3.
XX
PS Disclosure; Page 83-85; 94pp; English.
XX
CC The invention relates to a method of modulating angiogenesis and immune
CC response. Method involves administering to an individual a compound that
CC modulate the activity of signal transducer and activator of transcription
CC 3 (Stat3). Modulating angiogenesis is useful for treating or preventing
CC hypoxic or ischaemic condition or disorder which is the result of stroke,
CC ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,
CC tissue ischaemia in the lower extremities, infarction, trauma, vascular
CC occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,
CC chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,
CC epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest,
CC nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy
CC with neovascularisation. Suppressing an immune response is useful for
CC ameliorating a symptom of an autoimmune disease such as systemic lupus
CC erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,
CC Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,
CC mixed connective tissue disease, primary biliary cirrhosis, pernicious
CC anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,
CC gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,
CC idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's
CC disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune
CC infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and
CC dense deposit disease. The method is useful in preventing or treating
CC specific proliferative and oncogenic disease which includes sarcomas and
CC carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,
CC fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,
CC hypoproliferative disorders, physical trauma, lesions and wounds. The
CC method is also used in gene therapy. The present sequence is human Stat3
CC protein
XX
SQ Sequence 769 AA;

Query Match 98.7%; Score 1157; DB 5; Length 769;
Best Local Similarity 98.7%; Pred. No. 7.8e-98;
Matches 226; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358
```

RESULT 12
AAE22056
ID AAE22056 standard; protein; 769 AA.
XX
AC AAE22056;
XX
DT 25-JUL-2002 (first entry)
XX
DE Human protein related to angiogenesis regulation.
XX

```
KW   Human; signal transducer and activator of transcription 3; ischaemia;
KW   immune response; Stat3; coronary atherosclerosis; vascular occlusion;
KW   hypoxia; stroke; angiogenesis; myocardial infarction; hypoglycaemia;
KW   inflammation; chronic obstructive pulmonary disease; cardiac arrest;
KW   insulin dependent diabetes mellitus; emphysema; trauma; scleroderma;
KW   shock; chronic active hepatitis; adult respiratory distress syndrome;
KW   nitrogen necrosis; proliferative angiopathy; autoimmune thyroiditis;
KW   Sjogren's syndrome; multiple sclerosis; Addison's disease; epilepsy;
KW   polymyositis; rheumatoid arthritis; autoimmune infertility; anaemia;
KW   proliferative disease; Grave's disease; ulcerative colitis; sarcoma;
KW   carcinoma; degenerative disorder; gene therapy; growth deficiency;
KW   cirrhosis; hypoproliferative disorder; lesion.
XX
OS   Homo sapiens.
XX
PN   WO200220032-A1.
XX
PD   14-MAR-2002.
XX
PF   10-SEP-2001; 2001WO-US028254.
XX
PR   08-SEP-2000; 2000US-0231212P.
XX
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (UYSF-) UNIV SOUTH FLORIDA.
XX
PI   Yu H,  Pardoll D,  Jove R,  Dalton W;
XX
DR   WPI; 2002-362218/39.
XX
PT   Modulating angiogenesis and an immune response in an individual, for
PT   treating a hypoxic or ischemic condition, comprises administering a
PT   compound that modulates the activity of a signal transducer and activator
PT   of transcription 3.
XX
PS   Disclosure; Page 83-85; 94pp; English.
XX
CC   The invention relates to a method of modulating angiogenesis and immune
CC   response. Method involves administering to an individual a compound that
CC   modulate the activity of signal transducer and activator of transcription
CC   3 (Stat3). Modulating angiogenesis is useful for treating or preventing
CC   hypoxic or ischaemic condition or disorder which is the result of stroke,
CC   ischaemia, coronary atherosclerosis, myocardial infarction, inflammation,
CC   tissue ischaemia in the lower extremities, infarction, trauma, vascular
CC   occlusion, prenatal or postnatal oxygen deprivation, suffocation, shock,
CC   chronic obstructive pulmonary disease, choking, asphyxia, hypoglycaemia,
CC   epilepsy, emphysema, adult respiratory distress syndrome, cardiac arrest,
CC   nitrogen necrosis, proliferative angiopathy e.g. diabetic microangiopathy
CC   with neovascularisation. Suppressing an immune response is useful for
CC   ameliorating a symptom of an autoimmune disease such as systemic lupus
CC   erythematosus, multiple sclerosis, insulin dependent diabetes mellitus,
CC   Sjogren's syndrome, scleroderma, polymyositis, chronic active hepatitis,
CC   mixed connective tissue disease, primary biliary cirrhosis, pernicious
CC   anaemia, autoimmune thyroiditis, idiopathic Addison's disease, vitiligo,
CC   gluten-sensitive enteropathy, autoimmune neutropenia, myasthenia gravis,
CC   idiopathic thrombocytopenia purpura, Grave's disease, Goodpasture's
CC   disease, rheumatoid arthritis, cirrhosis, pemphigus vulgaris, autoimmune
CC   infertility, bullous pemphigoid, discoid lupus, ulcerative colitis and
CC   dense deposit disease. The method is useful in preventing or treating
CC   specific proliferative and oncogenic disease which includes sarcomas and
CC   carcinomas e.g., bladder carcinoma, colon carcinoma, chronic leukaemia,
CC   fibrosarcoma, liposarcoma, degenerative disorders, growth deficiency,
CC   hypoproliferative disorders, physical trauma, lesions and wounds. The
CC   method is also used in gene therapy. The present sequence is human
CC   protein related to angiogenesis regulation
XX
SQ   Sequence 769 AA;

  Query Match             98.7%;  Score 1157;  DB 5;  Length 769;
  Best Local Similarity   98.7%;  Pred. No. 7.8e-98;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 13
AAR82995
ID   AAR82995 standard; protein; 770 AA.
XX
AC   AAR82995;
XX
DT   25-MAR-1996  (first entry)
XX
DE   Mouse liver acute phase response factor.
XX
KW   Mouse; acute phase response factor; transcription factor; interleukin-6;
KW   signal transmission; liver; antibody; antisense; ribozyme;
KW   antiinflammatory; antitumor; hypotensive; therapy.
XX
OS   Mus musculus.
XX
PN   EP676469-A2.
XX
PD   11-OCT-1995.
XX
PF   29-MAR-1995;   95EP-00104670.
XX
PR   04-APR-1994;   94JP-00065825.
XX
PA   (KISH/) KISHIMOTO T.
XX
PI   Akira S,  Kishimoto T;
XX
DR   WPI; 1995-346089/45.
DR   N-PSDB; AAT05619.
XX
PT   New acute phase response factor - for developing inhibitory agents for
PT   treating diseases induced by cytokine(s) such as IL-6, e.g. inflammatory
PT   diseases.
XX
PS   Claim 10; Page 20-22; 31pp; English.
XX
CC   The sequence represents a mouse acute phase response factor (APRF), a
CC   transcription factor related to signal transmission of interleukin-6 (IL-
CC   6). The protein is encoded by a cDNA, isolated from a mouse liver cDNA
CC   library using a polymerase chain reaction product (amplified using
CC   primers derived from an IL-6-treated mouse liver peptide) as a probe.
CC   APRF-inhibitors, e.g. antibodies, antisense oligonucleotides or
CC   ribozymes, may be used to treat diseases induced by IL-6, e.g.
CC   inflammatory disease, leukemia, cancer, osteoclasia, pulmonary
CC   hypertension, etc
XX
SQ   Sequence 770 AA;

  Query Match             98.7%;  Score 1157;  DB 2;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 7.8e-98;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189
```

```
Qy          61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db         190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy         121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy         181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
               |||||||||||||||||||||||||||||||||||||||||||||||||
Db         310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 14
AAY03768
ID   AAY03768 standard; protein; 770 AA.
XX
AC   AAY03768;
XX
DT   11-JUN-1999  (first entry)
XX
DE   Human STAT3 allelic variant.
XX
KW   Signal Transducer and Activator of Transcription 3; STAT3; allele; IL-6;
KW   intracellular transcription factor; interleukin-6; medicament; variant;
KW   pharmaceutical; autoimmune disease; inflammatory; human.
XX
OS   Homo sapiens.
XX
PN   EP905234-A2.
XX
PD   31-MAR-1999.
XX
PF   18-FEB-1998;   98EP-00102774.
XX
PR   16-SEP-1997;   97EP-00116061.
XX
PA   (ISTF ) ARS APPLIED RES SYSTEMS HOLDING NV.
XX
PI   Serlupi-Crescenzi O,  Della Pietra L;
XX
DR   WPI; 1999-192664/17.
DR   N-PSDB; AAX29281.
XX
PT   New human Signal Transducer and Activator of Transcription 3 (STAT3)
PT   allelic variant useful for treatment of autoimmune and inflammatory
PT   disease.
XX
PS   Claim 2; Page 9-13; 32pp; English.
XX
CC   The present sequence represents a predominant allelic variant of human
CC   Signal Transducer and Activator of Transcription 3 (STAT3) protein, an
CC   intracellular transcription factor which mediates IL-6 signals. The
CC   encoding sequence differs from the original published human STAT3 gene
CC   sequence (Akira et al., 1994). Host cells comprising a recombinant STAT3
CC   DNA molecule can be used for the recombinant expression of the variant.
CC   STAT3 protein is useful as a medicament or pharmaceutical composition for
CC   treatment of autoimmune or inflammatory diseases
XX
SQ   Sequence 770 AA;

  Query Match             98.7%;  Score 1157;  DB 2;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 7.8e-98;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy           1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
               |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy          61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db         190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy         121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy         181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
               |||||||||||||||||||||||||||||||||||||||||||||||||
Db         310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 15
AAB12377
ID   AAB12377 standard; peptide; 770 AA.
XX
AC   AAB12377;
XX
DT   08-NOV-2000  (first entry)
XX
DE   N-terminal domain of murine STAT-3 protein.
XX
KW   STAT; signal transducer and activator of transcription; crystal;
KW   drug design; murine.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Region          4. .9
FT                   /label= Alpha helix 1
FT   Region          12. .21
FT                   /label= Alpha helix 2
FT   Region          19. .21
FT                   /label= 3(10) helix of alpha helix 2
FT   Region          28. .33
FT                   /label= Alpha helix 3
FT   Region          35. .40
FT                   /label= Alpha helix 4
FT   Region          43. .47
FT                   /label= Alpha helix 5
FT   Region          50. .73
FT                   /label= Alpha helix 6
FT   Region          77. .96
FT                   /label= Alpha helix 7
FT   Region          99. .119
FT                   /label= Alpha helix 8
XX
PN   US6087478-A.
XX
PD   11-JUL-2000.
XX
PF   23-JAN-1998;   98US-00012710.
XX
PR   23-JAN-1998;   98US-00012710.
XX
PA   (UYRQ ) UNIV ROCKEFELLER.
XX
PI   Vinkemeier U,  Moarefi I,  Darnell JE,  Kuriyan J;
XX
DR   WPI; 2000-505108/45.
XX
PT   New crystals of an N-terminal fragment of a signal transducer and
PT   activator of transcription that effectively diffracts x-rays, useful for
PT   drug screening and development.
XX
PS   Disclosure; Fig 1; 42pp; English.
XX
CC   The present invention relates to a crystal of an N-terminal fragment of a
CC   signal transducer and activator of transcription (STAT) protein. The
CC   crystal effectively diffracts X-rays, allowing the determination of the
CC   atomic coordinates of the N-terminal domain to a resolution of greater
CC   than 5.0 Angstroms. The present sequence is the N-terminal domain of the
CC   murine STAT 3 protein. The N-terminal domain enables STAT dimers to
CC   interact and bind DNA cooperatively, a mechanism important for gene
CC   activation. The crystals are useful in drug screening and development by
CC   selecting a potential drug by performing rational drug design with the 3-
```

```
CC   dimensional structure determined for the crystal
XX
SQ   Sequence 770 AA;

  Query Match             98.7%;  Score 1157;  DB 3;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 7.8e-98;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


Search completed: May 25, 2005, 17:36:39
Job time : 123.057 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:25:01 ; Search time 28.6534 Seconds
(without alignments)
596.600 Million cell updates/sec

Title: US-10-090-185-31
Perfect score: 1172
Sequence: 1 NHPTAAVVTEKQQMLEQHLQ.....IKTGVQFTTKVRLLVKFPEL 229

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1172	100.0	229	3	US-09-387-418A-31	Sequence 31, Appl
2	1168	99.7	229	3	US-09-387-418A-18	Sequence 18, Appl
3	1168	99.7	229	3	US-09-387-418A-28	Sequence 28, Appl
4	1168	99.7	252	3	US-09-387-418A-14	Sequence 14, Appl
5	1168	99.7	271	3	US-09-387-418A-9	Sequence 9, Appli
6	1168	99.7	770	1	US-08-369-796-12	Sequence 12, Appl
7	1168	99.7	770	2	US-08-852-091-12	Sequence 12, Appl
8	1168	99.7	770	2	US-08-820-754-12	Sequence 12, Appl
9	1168	99.7	770	3	US-08-956-652-12	Sequence 12, Appl
10	1168	99.7	770	3	US-08-956-869-12	Sequence 12, Appl
11	1168	99.7	770	3	US-08-948-547-12	Sequence 12, Appl
12	1168	99.7	770	3	US-09-364-970-3	Sequence 3, Appli
13	1168	99.7	770	3	US-09-364-970-5	Sequence 5, Appli
14	1168	99.7	770	3	US-08-956-653A-12	Sequence 12, Appl
15	1168	99.7	770	4	US-08-212-185-12	Sequence 12, Appl
16	1168	99.7	770	5	PCT-US95-17025-12	Sequence 12, Appl
17	1163	99.2	229	3	US-09-387-418A-30	Sequence 30, Appl
18	1157	98.7	770	1	US-08-416-581B-9	Sequence 9, Appli
19	1157	98.7	770	3	US-09-012-710-8	Sequence 8, Appli
20	1157	98.7	770	3	US-09-556-273-8	Sequence 8, Appli
21	1157	98.7	770	3	US-09-526-542-2	Sequence 2, Appli
22	1157	98.7	770	4	US-10-117-087-2	Sequence 2, Appli
23	1152	98.3	770	1	US-08-416-581B-1	Sequence 1, Appli
24	1152	98.3	770	1	US-08-416-581B-5	Sequence 5, Appli
25	1152	98.3	770	3	US-09-087-465-6	Sequence 6, Appli
26	1152	98.3	770	4	US-09-972-800A-6	Sequence 6, Appli
27	1152	98.3	771	1	US-08-276-099A-14	Sequence 14, Appl
28	1152	98.3	771	1	US-08-781-890-14	Sequence 14, Appl
29	1151	98.2	229	3	US-09-387-418A-29	Sequence 29, Appl
30	1090	93.0	213	3	US-09-387-418A-19	Sequence 19, Appl
31	1090	93.0	236	3	US-09-387-418A-15	Sequence 15, Appl
32	1044	89.1	223	3	US-09-387-418A-22	Sequence 22, Appl
33	845	72.1	185	3	US-09-387-418A-23	Sequence 23, Appl
34	777	66.3	176	3	US-09-387-418A-16	Sequence 16, Appl
35	653	55.7	128	3	US-09-387-418A-20	Sequence 20, Appl
36	601	51.3	143	3	US-09-387-418A-17	Sequence 17, Appl
37	571	48.7	129	3	US-09-387-418A-24	Sequence 24, Appl
38	569	48.5	268	3	US-09-387-418A-12	Sequence 12, Appl
39	569	48.5	582	4	US-09-430-806A-3	Sequence 3, Appli
40	569	48.5	712	1	US-08-369-796-6	Sequence 6, Appli
41	569	48.5	712	2	US-08-852-091-6	Sequence 6, Appli
42	569	48.5	712	2	US-08-820-754-6	Sequence 6, Appli
43	569	48.5	712	3	US-08-956-652-6	Sequence 6, Appli
44	569	48.5	712	3	US-08-956-869-6	Sequence 6, Appli
45	569	48.5	712	3	US-08-948-547-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-09-387-418A-31
; Sequence 31, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT: Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/09/387,418A
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-387-418A-31

Query Match 100.0%; Score 1172; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 2.7e-99;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
```

RESULT 2
US-09-387-418A-18
; Sequence 18, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H

```
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 18
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-18

  Query Match             99.7%;  Score 1168;  DB 3;  Length 229;
  Best Local Similarity   99.6%;  Pred. No. 6.3e-99;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 3
US-09-387-418A-28
; Sequence 28, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
;  APPLICANT:  Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 28
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-28

  Query Match             99.7%;  Score 1168;  DB 3;  Length 229;
  Best Local Similarity   99.6%;  Pred. No. 6.3e-99;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 4
US-09-387-418A-14
; Sequence 14, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
;  APPLICANT:  Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 252
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-14

  Query Match             99.7%;  Score 1168;  DB 3;  Length 252;
  Best Local Similarity   99.6%;  Pred. No. 7.1e-99;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         24 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 83

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 203

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 252


RESULT 5
US-09-387-418A-9
; Sequence 9, Application US/09387418A
; Patent No. 6391572
; GENERAL INFORMATION:
;  APPLICANT:  Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION:  METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE:  600-1-253
;  CURRENT APPLICATION NUMBER: US/09/387,418A
;  CURRENT FILING DATE:  1999-08-31
;  NUMBER OF SEQ ID NOS:  43
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 271
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-387-418A-9

  Query Match             99.7%;  Score 1168;  DB 3;  Length 271;
  Best Local Similarity   99.6%;  Pred. No. 7.8e-99;
```

```
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db         24 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 83

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 203

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 252


RESULT 6
US-08-369-796-12
; Sequence 12, Application US/08369796
; Patent No. 5716622
;  GENERAL INFORMATION:
;    APPLICANT:  James E. Darnell, Jr.
;    APPLICANT:  Zilong Wen
;    APPLICANT:  Curt M. Horvath
;    APPLICANT:  Zhong Zhong
;    TITLE OF INVENTION:  FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
;    TITLE OF INVENTION:  TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/369,796
;      FILING DATE:  06-JAN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-116
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-369-796-12

  Query Match             99.7%;  Score 1168;  DB 1;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 3e-98;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 7
US-08-852-091-12
; Sequence 12, Application US/08852091
; Patent No. 5883228
;  GENERAL INFORMATION:
;    APPLICANT:  James E. Darnell, Jr.
;    APPLICANT:  Zilong Wen
;    APPLICANT:  Curt M. Horvath
;    APPLICANT:  Zhong Zhong
;    TITLE OF INVENTION:  FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
;    TITLE OF INVENTION:  TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/852,091
;      FILING DATE:  06-MAY-1997
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/369,796
;      FILING DATE:  06-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-116
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-852-091-12

  Query Match             99.7%;  Score 1168;  DB 2;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 3e-98;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358
```

```
Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 8
US-08-820-754-12
; Sequence 12, Application US/08820754
; Patent No. 5976835
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/820,754
;      FILING DATE:  19-MAR-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,185
;      FILING DATE:  11-MAR-1994
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-820-754-12

  Query Match             99.7%;  Score 1168;  DB 2;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 3e-98;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189
```

```
Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 9
US-08-956-652-12
; Sequence 12, Application US/08956652
; Patent No. 6013475
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/956,652
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,185
;      FILING DATE:  11-MAR-1994
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-956-652-12

  Query Match             99.7%;  Score 1168;  DB 3;  Length 770;
```

```
  Best Local Similarity   99.6%;  Pred. No. 3e-98;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 10
US-08-956-869-12
; Sequence 12, Application US/08956869
; Patent No. 6030808
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/956,869
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/212,185
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-956-869-12

  Query Match             99.7%;  Score 1168;  DB 3;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 3e-98;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 11
US-08-948-547-12
; Sequence 12, Application US/08948547
; Patent No. 6124118
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/948,547
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/212,185
;      FILING DATE:  11-MAR-1994
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-948-547-12

  Query Match             99.7%;  Score 1168;  DB 3;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 3e-98;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 12
US-09-364-970-3
; Sequence 3, Application US/09364970
; Patent No. 6235873
; GENERAL INFORMATION:
;  APPLICANT: Bromberg, Jacqueline
;  TITLE OF INVENTION: CONSTITUTIVELY ACTIVE TRANSCRIPTION FACTORS AND THEIR
;  TITLE OF INVENTION: USES FOR IDENTIFYING MODULATORS OF ACTIVITY INCLUDING
;  TITLE OF INVENTION: DYSPROLIFERATIVE CELLULAR CHANGES
;  FILE REFERENCE: 600-1-252
;  CURRENT APPLICATION NUMBER: US/09/364,970
;  CURRENT FILING DATE: 1999-07-31
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-364-970-3

  Query Match             99.7%;  Score 1168;  DB 3;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 3e-98;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 13
US-09-364-970-5
; Sequence 5, Application US/09364970
; Patent No. 6235873
; GENERAL INFORMATION:
;  APPLICANT: Bromberg, Jacqueline
;  TITLE OF INVENTION: CONSTITUTIVELY ACTIVE TRANSCRIPTION FACTORS AND THEIR
;  TITLE OF INVENTION: USES FOR IDENTIFYING MODULATORS OF ACTIVITY INCLUDING
;  TITLE OF INVENTION: DYSPROLIFERATIVE CELLULAR CHANGES
;  FILE REFERENCE: 600-1-252
;  CURRENT APPLICATION NUMBER: US/09/364,970
;  CURRENT FILING DATE: 1999-07-31
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-364-970-5

  Query Match             99.7%;  Score 1168;  DB 3;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 3e-98;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 14
US-08-956-653A-12
; Sequence 12, Application US/08956653A
; Patent No. 6338949
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/956,653A
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,185
;      FILING DATE:  11-MAR-1994
```

```
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-195
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-956-653A-12

  Query Match             99.7%;  Score 1168;  DB 3;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 3e-98;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 15
US-08-212-185-12
; Sequence 12, Application US/08212185
; Patent No. 6605442
;  GENERAL INFORMATION:
;    APPLICANT:  Darnell Jr., James E.
;    APPLICANT:  Schindler, Christian W.
;    APPLICANT:  Fu, Xian-Yuan
;    APPLICANT:  Wen, Zilong
;    APPLICANT:  Zhong, Zhong
;    TITLE OF INVENTION:  RECEPTOR RECOGNITION FACTORS, PROTEIN
;    TITLE OF INVENTION:  SEQUENCES AND METHODS OF USE THEREOF
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/212,185
;      FILING DATE:  11-MAR-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/980,498
;      FILING DATE:  23-NOV-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/854,296
;      FILING DATE:  19-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO US93/02569
;      FILING DATE:  19-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/126,588
;      FILING DATE:  24-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-073 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201 487-5800
;      TELEFAX:  201 343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  770 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-212-185-12

  Query Match             99.7%;  Score 1168;  DB 4;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 3e-98;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


Search completed: May 25, 2005, 17:47:41
Job time : 29.6534 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:36:58 ; Search time 100.287 Seconds
(without alignments)
763.830 Million cell updates/sec

Title: US-10-090-185-31
Perfect score: 1172
Sequence: 1 NHPTAAVVTEKQQMLEQHLQ.....IKTGVQFTTKVRLLVKFPEL 229

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1434725 seqs, 334507595 residues

Total number of hits satisfying chosen parameters: 1434725

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1172	100.0	229	13	US-10-090-185-31	Sequence 31, Appl
2	1168	99.7	229	13	US-10-090-185-18	Sequence 18, Appl
3	1168	99.7	229	13	US-10-090-185-28	Sequence 28, Appl
4	1168	99.7	252	13	US-10-090-185-14	Sequence 14, Appl
5	1168	99.7	271	13	US-10-090-185-9	Sequence 9, Appli
6	1168	99.7	770	11	US-09-876-773-12	Sequence 12, Appl
7	1168	99.7	770	17	US-10-639-617-12	Sequence 12, Appl
8	1163	99.2	229	13	US-10-090-185-30	Sequence 30, Appl
9	1157	98.7	720	15	US-10-380-020-4	Sequence 4, Appli
10	1157	98.7	769	15	US-10-380-020-2	Sequence 2, Appli
11	1157	98.7	769	15	US-10-380-020-5	Sequence 5, Appli
12	1157	98.7	770	14	US-10-045-792-8	Sequence 8, Appli
13	1157	98.7	770	14	US-10-038-010-56	Sequence 56, Appl
14	1157	98.7	770	14	US-10-117-087-2	Sequence 2, Appli
15	1157	98.7	793	9	US-09-925-302-780	Sequence 780, App
16	1157	98.7	793	10	US-09-925-302-780	Sequence 780, App
17	1154	98.5	770	15	US-10-116-275-329	Sequence 329, App
18	1152	98.3	770	15	US-10-116-275-349	Sequence 349, App
19	1151	98.2	229	13	US-10-090-185-29	Sequence 29, Appl
20	1090	93.0	213	13	US-10-090-185-19	Sequence 19, Appl
21	1090	93.0	236	13	US-10-090-185-15	Sequence 15, Appl
22	1044	89.1	223	13	US-10-090-185-22	Sequence 22, Appl
23	845	72.1	185	13	US-10-090-185-23	Sequence 23, Appl
24	777	66.3	176	13	US-10-090-185-16	Sequence 16, Appl
25	653	55.7	128	13	US-10-090-185-20	Sequence 20, Appl
26	601	51.3	143	13	US-10-090-185-17	Sequence 17, Appl
27	573	48.9	749	9	US-09-833-205-4	Sequence 4, Appli
28	571	48.7	129	13	US-10-090-185-24	Sequence 24, Appl
29	569	48.5	268	13	US-10-090-185-12	Sequence 12, Appl
30	569	48.5	582	14	US-10-245-120-3	Sequence 3, Appli
31	569	48.5	712	11	US-09-876-773-6	Sequence 6, Appli
32	569	48.5	712	14	US-10-245-120-2	Sequence 2, Appli
33	569	48.5	712	17	US-10-639-617-6	Sequence 6, Appli
34	569	48.5	712	17	US-10-936-390-5	Sequence 5, Appli
35	569	48.5	750	9	US-09-833-205-2	Sequence 2, Appli
36	569	48.5	750	11	US-09-876-773-4	Sequence 4, Appli
37	569	48.5	750	14	US-10-245-120-1	Sequence 1, Appli
38	569	48.5	750	14	US-10-308-279-44	Sequence 44, Appl
39	569	48.5	750	16	US-10-755-889-352	Sequence 352, App
40	569	48.5	750	16	US-10-755-889-823	Sequence 823, App
41	569	48.5	750	17	US-10-492-043-19	Sequence 19, Appl
42	569	48.5	750	17	US-10-639-617-4	Sequence 4, Appli
43	569	48.5	786	9	US-09-925-297-550	Sequence 550, App
44	562	48.0	749	9	US-09-833-205-6	Sequence 6, Appli
45	562	48.0	749	11	US-09-876-773-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-10-090-185-31
; Sequence 31, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Xiaokui
; APPLICANT: Wrzeszczynska, Melissa H
; APPLICANT; Horvath, Curt M
; APPLICANT: Darnell Jr., James E
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
; TITLE OF INVENTION: INTERACTIONS
; FILE REFERENCE: 600-1-253
; CURRENT APPLICATION NUMBER: US/10/090,185
; CURRENT FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: 09/387,418
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-090-185-31

Query Match 100.0%; Score 1172; DB 13; Length 229;
Best Local Similarity 100.0%; Pred. No. 7.5e-95;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
```

```
Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 2
US-10-090-185-18
; Sequence 18, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-18

  Query Match             99.7%;  Score 1168;  DB 13;  Length 229;
  Best Local Similarity   99.6%;  Pred. No. 1.7e-94;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 3
US-10-090-185-28
; Sequence 28, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-28

  Query Match             99.7%;  Score 1168;  DB 13;  Length 229;
  Best Local Similarity   99.6%;  Pred. No. 1.7e-94;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 4
US-10-090-185-14
; Sequence 14, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
;   LENGTH: 252
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-14

  Query Match             99.7%;  Score 1168;  DB 13;  Length 252;
  Best Local Similarity   99.6%;  Pred. No. 1.9e-94;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db         24 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 83

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 203

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 252


RESULT 5
US-10-090-185-9
```

```
; Sequence 9, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR P
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 271
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-9

  Query Match             99.7%;  Score 1168;  DB 13;  Length 271;
  Best Local Similarity   99.6%;  Pred. No. 2.1e-94;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db         24 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 83

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 143

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 203

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        204 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 252


RESULT 6
US-09-876-773-12
; Sequence 12, Application US/09876773
; Publication No. US20040058318A1
;  GENERAL INFORMATION:
;       APPLICANT: Darnell Jr., James E.
;                  Schindler, Christian W.
;                  Fu, Xian-Yuan
;                  Wen, Zilong
;                  Zhong, Zhong
;       TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
;                           SEQUENCES AND METHODS OF USE THEREOF
;       NUMBER OF SEQUENCES: 25
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Klauber & Jackson
;            STREET: 411 Hackensack Avenue
;            CITY: Hackensack
;            STATE: New Jersey
;            COUNTRY: USA
;            ZIP: 07601
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/876,773
;            FILING DATE: 07-Jun-2001
;            CLASSIFICATION: 435
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/212,185
;            FILING DATE: 11-MAR-1994
;            APPLICATION NUMBER: US 07/980,498
;            FILING DATE: 23-NOV-1992
;            APPLICATION NUMBER: US 07/854,296
;            FILING DATE: 19-MAR-1992
;            APPLICATION NUMBER: WO US93/02569
;            FILING DATE: 19-MAR-1993
;            APPLICATION NUMBER: US 08/126,588
;            FILING DATE: 24-SEP-1993
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Jackson Esq., David A.
;            REGISTRATION NUMBER: 26,742
;            REFERENCE/DOCKET NUMBER: 600-1-073 CIP
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 201 487-5800
;            TELEFAX: 201 343-1684
;            TELEX: 133521
;  INFORMATION FOR SEQ ID NO: 12:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 770 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-876-773-12

  Query Match             99.7%;  Score 1168;  DB 11;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 7.9e-94;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 7
US-10-639-617-12
; Sequence 12, Application US/10639617
; Publication No. US20050079543A1
;  GENERAL INFORMATION:
;       APPLICANT: Darnell Jr., James E.
;                  Schindler, Christian W.
;                  Fu, Xian-Yuan
;                  Wen, Zilong
;                  Zhong, Zhong
;       TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
;                           SEQUENCES AND METHODS OF USE THEREOF
;       NUMBER OF SEQUENCES: 25
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Klauber & Jackson
;            STREET: 411 Hackensack Avenue
;            CITY: Hackensack
;            STATE: New Jersey
;            COUNTRY: USA
;            ZIP: 07601
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/639,617
;            FILING DATE: 12-Aug-2003
;            CLASSIFICATION: 435
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/212,185
;            FILING DATE: 11-MAR-1994
;            APPLICATION NUMBER: US 07/980,498
;            FILING DATE: 23-NOV-1992
;            APPLICATION NUMBER: US 07/854,296
;            FILING DATE: 19-MAR-1992
;            APPLICATION NUMBER: WO US93/02569
;            FILING DATE: 19-MAR-1993
;            APPLICATION NUMBER: US 08/126,588
;            FILING DATE: 24-SEP-1993
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Jackson Esq., David A.
;            REGISTRATION NUMBER: 26,742
;            REFERENCE/DOCKET NUMBER: 600-1-073 CIP
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 201 487-5800
;            TELEFAX: 201 343-1684
;            TELEX: 133521
;   INFORMATION FOR SEQ ID NO: 12:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 770 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-639-617-12

  Query Match             99.7%;  Score 1168;  DB 17;  Length 770;
  Best Local Similarity   99.6%;  Pred. No. 7.9e-94;
  Matches  228;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 8
US-10-090-185-30
; Sequence 30, Application US/10090185
; Publication No. US20020197647A1
; GENERAL INFORMATION:
;  APPLICANT: Zhang, Xiaokui
;  APPLICANT:  Wrzeszczynska, Melissa H
;  APPLICANT:  Horvath, Curt M
;  APPLICANT:  Darnell Jr., James E
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING MODULATORS OF TRANSCRIPTIONAL ACTIVATOR F
;  TITLE OF INVENTION:  INTERACTIONS
;  FILE REFERENCE: 600-1-253
;  CURRENT APPLICATION NUMBER: US/10/090,185
;  CURRENT FILING DATE:  2002-03-04
;  PRIOR APPLICATION NUMBER: 09/387,418
;  PRIOR FILING DATE: 1999-08-31
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-090-185-30

  Query Match             99.2%;  Score 1163;  DB 13;  Length 229;
  Best Local Similarity   99.1%;  Pred. No. 4.6e-94;
  Matches  227;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||| || ||||||||||||||||||||||||||||||||||||||
Db          1 NHPTAAVVTEKQQMLEQHAQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229


RESULT 9
US-10-380-020-4
; Sequence 4, Application US/10380020
; Publication No. US20040052762A1
; GENERAL INFORMATION:
;  APPLICANT: Yu, Hua
;  APPLICANT:  Pardoll, Drew
;  APPLICANT:  Jove, Richard
;  APPLICANT:  Dalton, Willian
;  TITLE OF INVENTION: Stat3 Agonists and Antagonists and Therapeutic Uses Thereof
;  FILE REFERENCE: 10873-009-999
;  CURRENT APPLICATION NUMBER: US/10/380,020
;  CURRENT FILING DATE:  2003-03-07
;  PRIOR APPLICATION NUMBER: 60/231,212
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
;   LENGTH: 720
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-380-020-4

  Query Match             98.7%;  Score 1157;  DB 15;  Length 720;
  Best Local Similarity   98.7%;  Pred. No. 6.7e-93;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 10
US-10-380-020-2
; Sequence 2, Application US/10380020
; Publication No. US20040052762A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Yu, Hua
;  APPLICANT:  Pardoll, Drew
;  APPLICANT:  Jove, Richard
;  APPLICANT:  Dalton, Willian
;  TITLE OF INVENTION: Stat3 Agonists and Antagonists and Therapeutic Uses Thereof
;  FILE REFERENCE: 10873-009-999
;  CURRENT APPLICATION NUMBER: US/10/380,020
;  CURRENT FILING DATE:  2003-03-07
;  PRIOR APPLICATION NUMBER: 60/231,212
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 769
;   TYPE: PRT<
;   ORGANISM: Homo Sapiens
US-10-380-020-2

  Query Match           98.7%;  Score 1157;  DB 15;  Length 769;
  Best Local Similarity   98.7%;  Pred. No. 7.3e-93;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 11
US-10-380-020-5
; Sequence 5, Application US/10380020
; Publication No. US20040052762A1
; GENERAL INFORMATION:
;  APPLICANT: Yu, Hua
;  APPLICANT:  Pardoll, Drew
;  APPLICANT:  Jove, Richard
;  APPLICANT:  Dalton, Willian
;  TITLE OF INVENTION: Stat3 Agonists and Antagonists and Therapeutic Uses Thereof
;  FILE REFERENCE: 10873-009-999
;  CURRENT APPLICATION NUMBER: US/10/380,020
;  CURRENT FILING DATE:  2003-03-07
;  PRIOR APPLICATION NUMBER: 60/231,212
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 769
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-380-020-5

  Query Match           98.7%;  Score 1157;  DB 15;  Length 769;
  Best Local Similarity   98.7%;  Pred. No. 7.3e-93;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 12
US-10-045-792-8
; Sequence 8, Application US/10045792
; Publication No. US20030003563A1
;  GENERAL INFORMATION:
;       APPLICANT: Vinkemeier, Uwe
;                  Moarefi, Ismail
;                  Darnell, Jr., James E.
;                  Kuriyan, John
;       TITLE OF INVENTION: A CRYSTAL OF THE N-TERMINAL DOMAIN OF A
;                           STAT PROTEIN AND METHODS OF USE THEREOF
;       NUMBER OF SEQUENCES: 13
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Klauber & Jackson
;            STREET: 411 Hackensack Avenue, 4th Floor
;            CITY: Hackensack
;            STATE: New Jersey
;            COUNTRY: USA
;            ZIP: 07601
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/045,792
;            FILING DATE: 19-Oct-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/012,710
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Jackson Esq., David A.
;            REGISTRATION NUMBER: 26,742
;            REFERENCE/DOCKET NUMBER: 600-1-194
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 201-487-5800
;            TELEFAX: 201-343-1684
;            TELEX: 133521
;  INFORMATION FOR SEQ ID NO: 8:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 770 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       HYPOTHETICAL: NO
;       SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-045-792-8

  Query Match           98.7%;  Score 1157;  DB 14;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 7.3e-93;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249
```

```
Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 13
US-10-038-010-56
; Sequence 56, Application US/10038010
; Publication No. US20030040089A1
; GENERAL INFORMATION:
;  APPLICANT: HYBRIGENICS
;  APPLICANT:  Pierre, Legrain
;  TITLE OF INVENTION: Protein-protein interactions in adipocyte cells
;  FILE REFERENCE: B4767A
;  CURRENT APPLICATION NUMBER: US/10/038,010
;  CURRENT FILING DATE:  2002-07-23
;  PRIOR APPLICATION NUMBER: US 60/259,377
;  PRIOR FILING DATE: 2001-01-02
;  NUMBER OF SEQ ID NOS: 67
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: STAT3 : Transcription factor
;   LOCATION: (1)..(770)
;   OTHER INFORMATION:
US-10-038-010-56

  Query Match             98.7%;  Score 1157;  DB 14;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 7.3e-93;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 14
US-10-117-087-2
; Sequence 2, Application US/10117087
; Publication No. US20030166854A1
; GENERAL INFORMATION:
;  APPLICANT: SERLUPI-CRESCENZI, Ottaviano
;  APPLICANT:  DELLA PIETRA, Linda
;  TITLE OF INVENTION: ALLELIC VARIANTS OF HUMAN STAT3
;  FILE REFERENCE: SERLUPI=2
;  CURRENT APPLICATION NUMBER: US/10/117,087
;  CURRENT FILING DATE:  2002-04-08
;  PRIOR APPLICATION NUMBER: US/09/526,542
;  PRIOR FILING DATE: 2000-03-19
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 770
;   TYPE: PRT
;   ORGANISM: Human
US-10-117-087-2

  Query Match             98.7%;  Score 1157;  DB 14;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 7.3e-93;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 15
US-09-925-302-780
; Sequence 780, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA104
;  CURRENT APPLICATION NUMBER: US/09/925,302
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05918
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 896
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 780
;   LENGTH: 793
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-925-302-780

  Query Match             98.7%;  Score 1157;  DB 9;  Length 793;
  Best Local Similarity   98.7%;  Pred. No. 7.6e-93;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        153 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 212

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        213 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 272

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 332

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 381


Search completed: May 25, 2005, 18:21:55
Job time : 101.62 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:24:07 ; Search time 21.6038 Seconds
(without alignments)
1019.898 Million cell updates/sec

Title: US-10-090-185-31
Perfect score: 1172
Sequence: 1 NHPTAAVVTEKQQMLEQHLQ.....IKTGVQFTTKVRLLVKFPEL 229

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1157	98.7	770	2	I49508	ISGF3 p91-related
2	1152	98.3	770	2	A54444	DNA-binding protei
3	523.5	44.7	739	2	A46159	interferon-depende
4	508	43.3	748	2	A56047	gamma-interferon a
5	322	27.5	851	2	A46160	interferon alpha-i
6	268.5	22.9	786	2	I49274	mammary gland fact
7	268.5	22.9	793	2	S54772	mammary gland fact
8	262.5	22.4	794	2	G02317	transcription acti
9	237	20.2	794	2	S55527	mammary gland fact
10	127	10.8	848	2	A54740	interleukin-4-indu
11	125.5	10.7	837	2	I57557	DNA-Binding Protei
12	122.5	10.5	1208	2	AE1947	chromosome segrega
13	122	10.4	533	2	G72593	hypothetical prote
14	122	10.4	978	2	A70387	conserved hypothet
15	120.5	10.3	1166	2	T27075	hypothetical prote
16	116.5	9.9	1509	1	A27224	myosin heavy chain
17	116.5	9.9	2094	2	S33124	tpr protein - huma
18	114.5	9.8	284	2	C64527	M protein - Helico
19	114.5	9.8	857	2	S33821	median body protei
20	113.5	9.7	924	2	S06117	myosin heavy chain
21	113	9.6	1999	1	S21801	myosin heavy chain
22	112.5	9.6	1690	2	T13030	microtubule bindin
23	112.5	9.6	2007	1	B43402	myosin heavy chain
24	112.5	9.6	2253	2	T30336	nuclear/mitotic ap
25	112	9.6	289	2	S51193	epimorphin - mouse
26	112	9.6	1976	2	A59252	myosin heavy chain
27	111	9.5	284	2	S24401	tropomyosin 2, ske
28	111	9.5	734	2	T27055	hypothetical prote
29	111	9.5	853	2	T51505	hypothetical prote
30	111	9.5	1164	2	T24806	hypothetical prote
31	111	9.5	1188	2	G83960	chromosome segrega
32	110.5	9.4	1738	2	T14867	interaptin - slime
33	109.5	9.3	434	2	T43448	hypothetical prote
34	109.5	9.3	464	2	H90279	microtubule bindin
35	109.5	9.3	727	2	AC1814	hypothetical prote
36	109.5	9.3	1048	1	BVECSC	exonuclease (EC 3.
37	109.5	9.3	1957	2	T38077	hypothetical coile
38	109.5	9.3	2442	2	T08621	centrosome associa
39	109	9.3	1959	1	A33977	myosin heavy chain
40	109	9.3	3187	2	JC5837	364K Golgi complex
41	108.5	9.3	946	2	S28061	SCP1 protein - rat
42	108.5	9.3	1940	2	A59287	myosin heavy chain
43	108.5	9.3	1992	2	A47297	myosin heavy chain
44	108	9.2	2677	2	A38194	desmoplakin I - hu
45	107.5	9.2	527	2	S33068	myosin heavy chain

ALIGNMENTS

RESULT 1
I49508
ISGF3 p91-related transcription factor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I49508; I49009
R;Akira, S.; Nishio, Y.; Inoue, M.; Wang, X.J.; Wei, S.; Matsusaka, T.; Yoshida, K.; Su
Cell 77, 63-71, 1994
A;Title: Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related tr
A;Reference number: A54444; MUID:94208062; PMID:7512451
A;Accession: I49508
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-770 <RES>
A;Cross-references: UNIPROT:P42227; GB:L29278; NID:g476715; PIDN:AAA37254.1; PID:g47671
R;Raz, R.; Durbin, J.E.; Levy, D.E.
J. Biol. Chem. 269, 24391-24395, 1994
A;Title: Acute phase response factor and additional members of the interferon-stimulate
A;Reference number: I49009; MUID:95014185; PMID:7523373
A;Accession: I49009
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-393,'M',395-700,702-770 <RE2>
A;Cross-references: EMBL:U08378; NID:g473889; PIDN:AAA56668.1; PID:g473890
C;Genetics:
A;Gene: APRF
C;Superfamily: human signal transducer and transcription activator STAT5A

Query Match 98.7%; Score 1157; DB 2; Length 770;
Best Local Similarity 98.7%; Pred. No. 2.7e-69;
Matches 226; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358
```

RESULT 2
A54444
DNA-binding protein APRF - human

```
C;Species: Homo sapiens (man)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 09-Jul-2004
C;Accession: A54444
R;Akira, S.; Nishio, Y.; Inoue, M.; Wang, X.J.; Wei, S.; Matsusaka, T.; Yoshida, K.; Sud
Cell 77, 63-71, 1994
A;Title: Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related tra
A;Reference number: A54444; MUID:94208062; PMID:7512451
A;Accession: A54444
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-770 <RES>
A;Cross-references: UNIPROT:P40763; GB:L29277; NID:g475788; PID:g475789
C;Genetics:
A;Gene: GDB:STAT3; APRF
A;Cross-references: GDB:358950
A;Map position: 17q21-17q21
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: DNA binding; transcription factor

  Query Match             98.3%;  Score 1152;  DB 2;  Length 770;
  Best Local Similarity   98.3%;  Pred. No. 5.9e-69;
  Matches  225;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELHQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 3
A46159
interferon-dependent positive-acting transcription factor ISGF-3 91K chain - human
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A46159
R;Schindler, C.; Fu, X.Y.; Improta, T.; Aebersold, R.; Darnell Jr., J.E.
Proc. Natl. Acad. Sci. U.S.A. 89, 7836-7839, 1992
A;Title: Proteins of transcription factor ISGF-3: one gene encodes the 91-and 84-kDa ISG
A;Reference number: A46159; MUID:92366557; PMID:1502203
A;Accession: A46159
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid; protein
A;Residues: 1-739 <SCH>
A;Cross-references: UNIPROT:P42224
A;Experimental source: HeLa cells
A;Note: sequence extracted from NCBI backbone (NCBIP:110818)
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             44.7%;  Score 523.5;  DB 2;  Length 739;
  Best Local Similarity   45.8%;  Pred. No. 2.6e-27;
  Matches  103;  Conservative   49;  Mismatches   60;  Indels   13;  Gaps    2;

Qy          5 AAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQ 64
              : |: :||: |:  ::: : :|  :| ::| :|:|||::||  |||  |    : ||  :
Db        132 STVMLDKQKELDSKVRNVKDKVMCIEHEIKSLEDLQDEYDFKCKTL--QNREHETNGVAK 189

Qy         65 SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEIACIG 124
              |  :|:   |::|   ||  |: :| ::  ||:  |  |  | ::|| :|||| : ||||
Db        190 SDQKQEQLLLKKMYLMLDNKRKEVVHKIIELLNVTELTQNALINDELVEWKRRQQSACIG 249

Qy        125 GPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERIVELF 184
              |||| |||:|:           | |||:||||||:|| :|: ||| ::: :| :|   ||

Db        250 GPPNACLDQLQ----------QVRQQLKKLEELEQKYTYEHDPITKNKQVLWDRTFSLF 298

Qy        185 RNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              : |::|:|||||||||| || ||||:||||||| |:|||||  ||
Db        299 QQLIQSSFVVERQPCMPTHPQRPLVLKTGVQFTVKLRLLVKLQEL 343


RESULT 4
A56047
gamma-interferon activation site-binding protein Stat4 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 09-Jul-2004
C;Accession: A56047
R;Yamamoto, K.; Quelle, F.W.; Thierfelder, W.E.; Kreider, B.L.; Gilbert, D.J.; Jenkins,
Mol. Cell. Biol. 14, 4342-4349, 1994
A;Title: Stat4, a novel gamma interferon activation site-binding protein expressed in e
A;Reference number: A56047; MUID:94277038; PMID:8007943
A;Accession: A56047
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-748 <YAM>
A;Cross-references: UNIPROT:P42228; GB:U09351; NID:g509502; PIDN:AAA19692.1; PID:g50950
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: DNA binding; phosphoprotein

  Query Match             43.3%;  Score 508;  DB 2;  Length 748;
  Best Local Similarity   45.5%;  Pred. No. 2.8e-26;
  Matches  102;  Conservative   48;  Mismatches   66;  Indels    8;  Gaps    2;

Qy          8 VTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKS--QGDMQDLNGNNQS 65
              |:|:|: :|  :   :  ||  ||  | :|:|||:||: |||:::  |||       |
Db        136 VSERQRNVEHKVSAIKNSVQMTEQDTKYLEDLQDEFDYRYKTIQTMDQGD------KNSI 189

Qy         66 VTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEIACIGG 125
              :  |::  |::|| :||  |:  :|::  :::  : :  ::  ||| |||:| :||||||
Db        190 LVNQEVLTLQEMLNSLDFKRKEALSKMTQIVNETDLLMNSMLLEELQDWKKRQQIACIGG 249

Qy        126 PPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERIVELFR 185
              | :  ||:|:|  | ||||  | |||::||:|   |::|:|||    |  | ||   |
Db        250 PLHNGLDQLQNCFTLLAESLFQLRQQLEKLQEQSTKMTYEGDPIPAQRAHLLERATFLIY 309

Qy        186 NLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              || |::|||||||||| || ||:|:|| :||| |:|||:| |||
Db        310 NLFKNSFVVERQPCMPTHPQRPMVLKTLIQFTVKLRLLIKLPEL 353


RESULT 5
A46160
interferon alpha-induced transcription activator ISGF-3, 113K chain - human
N;Alternate names: stat2 protein
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A46160; S71908; S53873
R;Fu, X.Y.; Schindler, C.; Improta, T.; Aebersold, R.; Darnell Jr., J.E.
Proc. Natl. Acad. Sci. U.S.A. 89, 7840-7843, 1992
A;Title: The proteins of ISGF-3, the interferon alpha-induced transcriptional activator
A;Reference number: A46160; MUID:92366558; PMID:1502204
A;Accession: A46160
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA; protein
A;Residues: 1-851 <FU1>
A;Cross-references: UNIPROT:P52630
A;Note: sequence extracted from NCBI backbone (NCBIP:110820)
R;Yan, R.; Qureshi, S.; Zhong, Z.; Wen, Z.; Darnell, J.E.
submitted to the EMBL Data Library, December 1994
A;Reference number: S71908
A;Accession: S71908
A;Molecule type: DNA
A;Residues: 1-851 <YAN>
A;Cross-references: EMBL:U18671; NID:g1293919; PIDN:AAA98760.1; PID:g1293920
R;Yan, R.; Qureshi, S.; Zhong, Z.; Wen, Z.; Darnell Jr., J.E.
Nucleic Acids Res. 23, 459-463, 1995
A;Title: The genomic structure of the STAT genes: multiple exons in coincident sites in
```

```
A;Reference number: S53873; MUID:95192056; PMID:7885841
A;Accession: S53873
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-196;392-591;684-730 <YAW>
A;Cross-references: EMBL:U18671
C;Genetics:
A;Gene: stat2
A;Introns: 44/2; 95/3; 127/3; 157/3; 183/1; 211/3; 261/2; 314/2; 345/2; 365/2; 372/2; 40
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: signal transduction; transcription regulation

  Query Match             27.5%;  Score 322;  DB 2;  Length 851;
  Best Local Similarity   33.5%;  Pred. No. 6.9e-14;
  Matches   74;  Conservative   51;  Mismatches   90;  Indels    6;  Gaps    4;

Qy          8 VTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVT 67
              |  :|  :|  : | |  :: | : :  ::: || | | || ::::|    |: :
Db        138 VESQQHEIESRILDLRAMMEKLVKSISQLKDQQDVFCFRYK-IQAKGKTPSLDPH--QTK 194

Qy         68 RQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEIACIGGPP 127
               ||:  |:: |  ||: |: ::     ||  :  : : |   :| :|| : : |||  |
Db        195 EQKI--LQETLNELDKRRKEVLDASKALLGRLTTLIELLL-PKLEEWKAQQQKACIRAPI 251

Qy        128 NICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERIVELFRNL 187
              :  |::|| | |: |:     || :|:|: |   |||: ||: :   :   :: || : |
Db        252 DHGLEQLETWFTAGAKLLFHLRQLLKELKGLSCLVSYQDDPLTKGVDLRNAQVTELLQRL 311

Qy        188 MKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPE 228
              :  ||||| |||||  | |||::||| :|| : ||||:  |
Db        312 LHRAFVVETQPCMPQTPHRPLILKTGSKFTVRTRLLVRLQE 352


RESULT 6
I49274
mammary gland factor - mouse
N;Alternate names: STAT5 protein homolog p80
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I49274; S54773; S54727
R;Liu, X.; Robinson, G.W.; Gouilleux, F.; Groner, B.; Hennighausen, L.
Proc. Natl. Acad. Sci. U.S.A. 92, 8831-8835, 1995
A;Title: Cloning and expression of Stat5 and an additional homologue (Stat5b) involved i
A;Reference number: I49273; MUID:96004632; PMID:7568026
A;Accession: I49274
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-786 <RES>
A;Cross-references: UNIPROT:P42232; UNIPROT:Q9JKM1; EMBL:U21110; NID:g747973; PIDN:AAC52
R;Mui, A.L.F.; Wakao, H.; O'Farrell, A.M.; Harada, N.; Miyajima, A.
EMBO J. 14, 1166-1175, 1995
A;Title: Interleukin-3, granulocyte-macrophage colony stimulating factor and interleukin
A;Reference number: S54772; MUID:95237198; PMID:7720707
A;Accession: S54773
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-432,'E',434-786 <MUI>
A;Cross-references: EMBL:Z48539; NID:g758635; PIDN:CAA88420.1; PID:g758636
R;Azam, M.; Erdjument-Bromage, H.; Kreider, B.L.; Xia, M.; Quelle, F.; Basu, R.; Saris,
EMBO J. 14, 1402-1411, 1995
A;Title: Interleukin-3 signals through multiple isoforms of Stat5.
A;Reference number: S54725; MUID:95246733; PMID:7537213
A;Accession: S54727
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-432,'E',434-786 <AZA>
C;Genetics:
A;Gene: Stat5b
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             22.9%;  Score 268.5;  DB 2;  Length 786;
  Best Local Similarity   30.6%;  Pred. No. 2.2e-10;
  Matches   71;  Conservative   41;  Mismatches   99;  Indels   21;  Gaps    5;

Qy          5 AAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNY-KTLKSQGDMQDLNGNN 63
              |  :::|   : |  :: |   || | ::| ::  |: |   | ::|: |      |   |
Db        134 ADAMSQKHLQINQTFEELRLITQDTENELKKLQQTQEYFIIQYQESLRIQAQFAQLGQLN 193

Qy         64 -------QSVTRQKMQQLEQML----TALDQMRRSIVSELAGLLSAMEYVQKTLTDEELA 112
                     ::  :||   ||  |       | | |  :  :    |  :   |  : |:||
Db        194 PQERMSRETALQQKQVSLETWLQREAQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELI 253

Qy        113 DWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQH 172
              |||| ::|  ||||   || |::|   |||   | ||||:: | | |::   | |: :
Db        254 QWKRRQQLAGNGGPPEGSLDVLQSWCEKLAEIIWQNRQQIRRAEHLCQQLPIPG-PVEEM 312

Qy        173 RPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLV 224
                :   | ::   |: | |::|:||        | |:||  :|   |||||
Db        313 LAEVNATITDIISALVTSTFIIEKQP--------PQVLKTQTKFAATVRLLV 356


RESULT 7
S54772
mammary gland factor - mouse
N;Alternate names: stat5 protein
C;Species: Mus musculus (house mouse)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S54772; I49273
R;Mui, A.L.F.; Wakao, H.; O'Farrell, A.M.; Harada, N.; Miyajima, A.
EMBO J. 14, 1166-1175, 1995
A;Title: Interleukin-3, granulocyte-macrophage colony stimulating factor and interleuki
A;Reference number: S54772; MUID:95237198; PMID:7720707
A;Accession: S54772
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-793 <MUI>
A;Cross-references: UNIPROT:P42230; UNIPROT:Q9JIA0; EMBL:Z48538; NID:g758633; PIDN:CAA8
R;Liu, X.; Robinson, G.W.; Gouilleux, F.; Groner, B.; Hennighausen, L.
Proc. Natl. Acad. Sci. U.S.A. 92, 8831-8835, 1995
A;Title: Cloning and expression of Stat5 and an additional homologue (Stat5b) involved
A;Reference number: I49273; MUID:96004632; PMID:7568026
A;Accession: I49273
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-793 <RES>
A;Cross-references: EMBL:U21103; NID:g747971; PIDN:AAA80590.1; PID:g747972
C;Genetics:
A;Gene: Stat5a
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             22.9%;  Score 268.5;  DB 2;  Length 793;
  Best Local Similarity   30.6%;  Pred. No. 2.2e-10;
  Matches   70;  Conservative   41;  Mismatches   97;  Indels   21;  Gaps    5;

Qy          8 VTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNY-KTLKSQGDMQDLNGNN--- 63
              :::|   : |  :: |   || | ::| ::  |: |   | ::|: |      |   |
Db        137 MSQKHLQINQRFEELRLITQDTENELKKLQQTQEYFIIQYQESLRIQAQFAQLGQLNPQE 196

Qy         64 ----QSVTRQKMQQLEQML----TALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWK 115
                   ::  :||   ||  |       | | |  :  :    |  :   |  : |:||  ||
Db        197 RMSRETALQQKQVSLETWLQREAQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELIQWK 256

Qy        116 RRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPM 175
              || ::|  ||||   || |::|   |||   | ||||:: | | |::   | |: :
Db        257 RRQQLAGNGGPPEGSLDVLQSWCEKLAEIIWQNRQQIRRAEHLCQQLPIPG-PVEEMLAE 315

Qy        176 LEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLV 224
              :   | ::   |: | |::|:||        | |:||  :|   |||||
Db        316 VNATITDIISALVTSTFIIEKQP--------PQVLKTQTKFAATVRLLV 356


RESULT 8
G02317
transcription activator stat5A - human
C;Species: Homo sapiens (man)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 09-Jul-2004
```

```
C;Accession: G02317
R;Lin, J.
submitted to the EMBL Data Library, December 1995
A;Reference number: H01043
A;Accession: G02317
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-794 <LIN>
A;Cross-references: UNIPROT:P42229; EMBL:U43185; NID:g1151169; PIDN:AAB06589.1; PID:g115
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             22.4%;  Score 262.5;  DB 2;  Length 794;
  Best Local Similarity   30.1%;  Pred. No. 5.6e-10;
  Matches   69;  Conservative   42;  Mismatches   97;  Indels   21;  Gaps    5;

Qy          8 VTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNY-KTLKSQGDMQDL------- 59
              :::|   : |  :: |   || | ::| ::  |: |   | ::|: |     |
Db        137 MSQKHLQINQTFEELRLVTQDTENELKKLQQTQEYFIIQYQESLRIQAQFAQLAQLSPQE 196

Qy         60 NGNNQSVTRQKMQQLEQML----TALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWK 115
               : ::  :||   ||  |       | | |  :  :    |  :   |  : |:||  ||
Db        197 RLSRETALQQKQVSLEAWLQREAQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELIQWK 256

Qy        116 RRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPM 175
              || ::|  ||||   || |::|   |||   | ||||:: | | |::   | |: :
Db        257 RRQQLAGNGGPPEGSLDVLQSWCEKLAEIIWQNRQQIRRAEHLCQQLPIPG-PVEEMLAE 315

Qy        176 LEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLV 224
              :   | ::   |: | |::|:||       | |:||  :|   |||||
Db        316 VNATITDIISALVTSTFIIEKQP--------PQVLKTQTKFAATVRLLV 356


RESULT 9
S55527
mammary gland factor - sheep
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S55527; S44353
R;Wakao, H.; Gouilleux, F.; Groner, B.
EMBO J. 14, 854-855, 1995
A;Title: Corrigenda. Mammary gland factor (MGF) is a novel member of the cytokine regula
A;Reference number: S55527; MUID:95188889; PMID:7882987
A;Accession: S55527
A;Molecule type: mRNA
A;Residues: 1-794 <WAK>
A;Cross-references: UNIPROT:P42231; EMBL:X78428; NID:g602354; PIDN:CAA55191.1; PID:g6023
A;Note: this is a revision to the sequence from reference S44353
R;Wakao, H.; Gouilleux, F.; Groner, B.
EMBO J. 13, 2182-2191, 1994
A;Title: Mammary gland factor (MGF) is a novel member of the cytokine regulated transcri
A;Reference number: S44353; MUID:94244619; PMID:7514531
A;Accession: S44353
A;Molecule type: mRNA
A;Residues: 1-716,'RHLHGPGSLPSR',729,'P',731,'ASL' <WAW>
A;Cross-references: EMBL:X78428
A;Note: this sequence has been revised in reference S55527
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             20.2%;  Score 237;  DB 2;  Length 794;
  Best Local Similarity   29.3%;  Pred. No. 2.8e-08;
  Matches   67;  Conservative   41;  Mismatches   99;  Indels   22;  Gaps    6;

Qy          8 VTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNY-KTLKSQGDMQDLNGNN--- 63
              :::|   : |  :: |   || | ::| ::  |: |   | ::|: |     |   |
Db        138 MSQKHLQINQTFEELRLVTQDTENELKKLQQTQEYFIIQYQESLRIQAQFAQLAQLNPQE 197

Qy         64 ----QSVTRQKMQQLEQML----TALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWK 115
                 ::  :||   ||  |       | | |  :  :    |  :   |  : |:||  ||
Db        198 RLSRETALQQKQVSLEAWLQREAQTLQQYRVELAEKHQKTLQLLRKQQTIILDDELIQWK 257

Qy        116 RRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPM 175
              || :   :  ||   || |::|   |||   | ||||:: | | |::   | |: :
Db        258 RRHDWRGMEAPPR-SLDVLQSWCEKLAEIIWQNRQQIRRAEHLCQQLPIPG-PVEEMLAE 315

Qy        176 LEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLV 224
              :   | ::   |: | |::|:||       | |:||  :|   |||||
Db        316 VNATITDIISALVTSTFIIEKQP--------PQVLKTQTKFAATVRLLV 356


RESULT 10
A54740
interleukin-4-induced transcription factor stat - human
C;Species: Homo sapiens (man)
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004
C;Accession: A54740
R;Hou, J.; Schindler, U.; Henzel, W.J.; Ho, T.C.; Brasseur, M.; McKnight, S.L.
Science 265, 1701-1706, 1994
A;Title: An interleukin-4-induced transcription factor: IL-4 stat.
A;Reference number: A54740; MUID:94367369; PMID:8085155
A;Accession: A54740
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-848 <HOU>
A;Cross-references: UNIPROT:P42226
C;Superfamily: human signal transducer and transcription activator STAT5A
C;Keywords: DNA binding; transcription regulation

  Query Match             10.8%;  Score 127;  DB 2;  Length 848;
  Best Local Similarity   24.7%;  Pred. No. 0.58;
  Matches   59;  Conservative   37;  Mismatches   93;  Indels   50;  Gaps    9;

Qy          3 PTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGN 62
              |   | |        | ||  :| |       |:   :|   | :  : || :  :: |
Db         92 PLKLVAT-----FRQILQGKKKAV------MEQFRHLPMPFHWKQEELKFKTGLRRLQHR 140

Qy         63 NQSV--TRQKMQQ--------LEQML------TALDQMRRSIVSELAGLLSAMEYVQKTL 106
                 :  |: :|:        |  ::       |   :    :: |  | | |     | |
Db        141 VGEIHLLREALQKGAEAGQVSLHSLIETPANGTGPSEALAMLLQETTGELEA----AKAL 196

Qy        107 TDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKG 166
                : :  |||: ::|  | |           :  |||  | :    :    :|||:|   |
Db        197 VLKRIQIWKRQQQLAGNGAP----------FEESLAPLQERCESLVDIYSQLQQEVGAAG 246

Qy        167 DPI-VQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLV 224
                :  : |  |  |: |: | |: | |:||:||         | |:||  :|    || |:
Db        247 GELEPKTRASLTGRLDEVLRTLVTSCFLVEKQP--------PQVLKTQTKFQAGVRFLL 297


RESULT 11
I57557
DNA-Binding Protein and transcription factor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C;Accession: I57557
R;Quelle, F.W.; Shimoda, K.; Thierfelder, W.; Fischer, C.L.; Kim, A.; Ruben, S.M.; Clev
Mol. Cell. Biol. 15, 3336-3343, 1995
A;Title: Cloning of murine Stat6 and human Stat6, Stat proteins that are tyrosine phosp
A;Reference number: I57557; MUID:95280934; PMID:7760829
A;Accession: I57557
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-837 <RES>
A;Cross-references: UNIPROT:P52633; GB:L47650; NID:g1008876; PIDN:AAA79006.1; PID:g1008
C;Genetics:
A;Gene: STAT6
C;Superfamily: human signal transducer and transcription activator STAT5A

  Query Match             10.7%;  Score 125.5;  DB 2;  Length 837;
  Best Local Similarity   26.2%;  Pred. No. 0.72;
  Matches   61;  Conservative   31;  Mismatches   82;  Indels   59;  Gaps   10;

Qy         10 EKQQMLEQ--HLQDARKRVQD----------LEQKMKVVENLQDDFDFNYKTLKSQGDMQ 57
              ||: ::|:  ||      | |:          |  :::    |::       ||    | :|
Db        106 EKKAVIEEFRHLPGPFHRKQEELKFTTPLGRLHHRVRETRLLRESLHLGPKT--GQVSLQ 163

Qy         58 D-----LNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELA 112
```

```
Db        164 NLIDPPLNGPGPS---------EDLPTILQ-----------GTVGDLETTQ-PLVLLRIQ 202

Qy        113 DWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLE-ELQQKVSYKGDPIVQ 171

Db        203 IWKRQQQLAGNGTPFEESLAGLQERCESLVEIYSQLHQEIGAASGELEPKT--------- 253

Qy        172 HRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLV 224

Db        254 -RASLISRLDEVLRTLVTSSFLVEKQP--------PQVLKTQTKFQAGVRFLL 297


RESULT 12
AE1947
chromosome segregation protein [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AE1947
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AE1947
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1208 <KUR>
A;Cross-references: UNIPROT:Q8YXT3; GB:BA000019; PIDN:BAB73085.1; PID:g17130474; GSPDB:G
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr1128
C;Superfamily: chromosome segregation protein SMC1

  Query Match             10.5%;  Score 122.5;  DB 2;  Length 1208;
  Best Local Similarity   25.2%;  Pred. No. 1.7;
  Matches   58;  Conservative   47;  Mismatches   76;  Indels   49;  Gaps   11;

Qy          5 AAVVTEKQQMLEQH--LQDARKRVQDLE-QKMKVVENLQDDFD--FNYKTLK-------S 52

Db        835 ATIKTQEQQIQQRETAFREAEQRLKNLENQQQRLQERIQEAQQRITEYQTQQTTCTEAIN 894

Qy         53 QGDMQDLNGNNQ-SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEEL 111

Db        895 RVSQQTTTINAQITQTRAKLSELEQHLGAEKQKRDTIEQEVRSHLLRQQQLEWEIQKLEE 954

Qy        112 ADWKRRPEIACIGG----------------PPNICLDRLENWITSLAESQLQTRQQI--- 152

Db        955 TQLKRREDLTALQSQLQELVPELPNPLPEVPDKVDLEELQKELRSLAK-RLQAMEPVNML 1013

Qy        153 ---------KKLEELQQKV-SYKGDPIVQHRPMLEERIVELFRNLMKSAF 192

Db       1014 ALEEYERTQKRLEELSQKLQTLEGE-----RTELLLRI-ENFTTLRQIAF 1057


RESULT 13
G72593
hypothetical protein APE1216 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: G72593
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: G72593
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-533 <KAW>
A;Cross-references: UNIPROT:Q9YCP2; DDBJ:AP000061; NID:g5104821; PIDN:BAA80205.1; PID:d1
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1216

  Query Match             10.4%;  Score 122;  DB 2;  Length 533;
  Best Local Similarity   21.2%;  Pred. No. 0.73;
  Matches   42;  Conservative   48;  Mismatches   66;  Indels   42;  Gaps    7;

Qy          8 VTEKQQMLEQHLQDARKRVQDLEQKMKVVEN----LQDDF---------------DFNYK 48

Db        319 MSQQLQALAEDLESLSSRVEDLEARVGSVEDRLSQAEEDIDSLTTSLDSLRTELEDLSTR 378

Qy         49 TLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTD 108

Db        379 LAEAQASLEDLNTRLDQVA-STLQQLQQRLATAEESLQALTEDLASLQAEVETLQQSIVE 437

Qy        109 EELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQ----TRQQIKKLEELQQKVSY 164

Db        438 IDRRLGQLRSTVDAV---------RLE--VESLGEKLVQAEEKNQRQDASIEDFQSQIQ- 485

Qy        165 KGDPIVQHRPMLEERIVE 182

Db        486 ------ELRTQLDEKTRE 497


RESULT 14
A70387
conserved hypothetical protein aq_1006 - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C;Accession: A70387
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; C
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: A70387
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-978 <AQF>
A;Cross-references: UNIPROT:O67124; GB:AE000718; NID:g2983504; PIDN:AAC07092.1; PID:g29
A;Experimental source: strain VF5
C;Genetics:
A;Gene: aq_1006
C;Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

  Query Match             10.4%;  Score 122;  DB 2;  Length 978;
  Best Local Similarity   23.3%;  Pred. No. 1.5;
  Matches   49;  Conservative   38;  Mismatches   81;  Indels   42;  Gaps    6;

Qy         10 EKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNGNNQSVTRQ 69

Db        234 EEKDSLERELSQVVTKLKELENLEKEVEKLREKLEFSRKVAP------------YVPIAK 281

Qy         70 KMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLT---------------DEELAD- 113

Db        282 RIEEIDKKLTELKVRKNKLTKELAVLKDELSFAQEELNRIEAEKEKFKEEKEREKELEHR 341

Qy        114 WKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHR 173

Db        342 LKKLQEIKEI-------LKELSQLSSSLKEKEREYEQAKQEFEDLSERVE-KGKKLVAET 393

Qy        174 PMLEERIVELFRN------LMKSAFVVERQ 197

Db        394 EEKLEKIKELFSEEEYTSLKMKERLLVELQ 423


RESULT 15
T27075
hypothetical protein Y51A2D.16 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T27075
R;McMurray, A.
submitted to the EMBL Data Library, January 1998
```

```
A;Reference number: Z20307
A;Accession: T27075
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1166 <WIL>
A;Cross-references: EMBL:AL021497; PIDN:CAA16403.1; GSPDB:GN00023; CESP:Y51A2D.16
A;Experimental source: clone Y51A2D
C;Genetics:
A;Gene: CESP:Y51A2D.16
A;Map position: 5
A;Introns: 17/3; 45/1; 76/2; 156/3; 240/2; 288/2; 594/2; 657/3; 756/3; 798/2; 900/3; 948

  Query Match            10.3%;  Score 120.5;  DB 2;  Length 1166;
  Best Local Similarity   21.4%;  Pred. No. 2.3;
  Matches   52;  Conservative   46;  Mismatches   72;  Indels   73;  Gaps    9;

Qy          1 NHPTAAVVTEKQ----QMLEQHLQD-------------ARKRVQDLEQKMKVVENLQDD 42
              :| | :: :||     ::||  |||               ||::| :| :  : ||::: 
Db        430 SHVTRSLSSEKNTLKAKLLE--LQDQVEAQTLELLNQKNCGKRLEDRDQMISNLHNLKNE 487

Qy         43 FDFNYKTLKSQGDMQDLNGNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEY- 101
               : : || ::| :::           :|:|:| : |      |  ::  :  | : :
Db        488 LENDLKTCQTQLELES----------KKLQRLREDLVLEKSRRADLIGRIHSLCTTLSLN 537

Qy        102 ---VQKTLTDEELAD-----------WKRRPEIACIGGPPNI---------------- 129
                  :|   |:|| |             ||  :   | |   |
Db        538 GANFEKINNDDELIDNIDDIMMNALVAVKRERDDLRIQGNQQIQELHDLKRDIEKLRRSE 597

Qy        130 --CLDRLENWITSLAESQLQTRQQI----KKLEELQQKVSYKGDPIVQHRPMLEERIVEL 183
                 |:  :: :  |    : |::|:    :|| ||  ::| | | |     |::  | ||
Db        598 SESLNESDDRVRELTRENMHTKEQVFMLQEKLRELNLELSTKNDEI----DMVKASIEEL 653

Qy        184 FRN 186
               ||
Db        654 NRN 656


Search completed: May 25, 2005, 17:45:30
Job time : 22.6038 secs
```

OM protein - protein search, using sw model

Run on: May 25, 2005, 17:15:30 ; Search time 95.284 Seconds
(without alignments)
1230.701 Million cell updates/sec

Title: US-10-090-185-31
Perfect score: 1172
Sequence: 1 NHPTAAVVTEKQQMLEQHLQ.....IKTGVQFTTKVRLLVKFPEL 229

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1157	98.7	770	1	STA3_HUMAN	P40763 homo sapien
2	1157	98.7	770	1	STA3_MOUSE	P42227 mus musculu
3	1154	98.5	770	1	STA3_RAT	P52631 rattus norv
4	1150	98.1	770	1	STA3_BOVIN	P61635 bos taurus
5	1132	96.6	771	2	Q6DV79	Q6dv79 gallus gall
6	1099	93.8	769	2	Q9PVX8	Q9pvx8 xenopus lae
7	1096	93.5	766	2	Q7ZXK3	Q7zxk3 xenopus lae
8	1011	86.3	414	2	Q7ZTS5	Q7zts5 brachydanio
9	1011	86.3	765	2	Q6DVF3	Q6dvf3 oryzias lat
10	1011	86.3	785	2	Q6GUE7	Q6gue7 oryzias lat
11	1011	86.3	786	2	Q6NV46	Q6nv46 brachydanio
12	1011	86.3	806	2	O93599	O93599 brachydanio
13	987	84.2	764	2	Q90Y16	Q90y16 tetraodon f
14	977	83.4	767	2	O13133	O13133 oncorhynchu
15	656	56.0	163	2	Q9N145	Q9n145 macaca mula
16	575	49.1	751	2	Q8JGN0	Q8jgn0 xenopus lae
17	573	48.9	712	2	Q6P6Q7	Q6p6q7 rattus norv
18	573	48.9	749	2	Q9QXK0	Q9qxk0 rattus norv
19	571	48.7	749	2	Q8C497	Q8c497 mus musculu
20	570	48.6	712	2	Q99K94	Q99k94 mus musculu
21	570	48.6	749	2	Q8C3V4	Q8c3v4 mus musculu
22	570	48.6	749	2	Q9D323	Q9d323 mus musculu
23	570	48.6	755	2	Q8C8M3	Q8c8m3 mus musculu
24	569	48.5	750	1	STA1_HUMAN	P42224 homo sapien
25	569	48.5	750	2	Q68D00	Q68d00 homo sapien
26	569	48.5	757	2	Q764M5	Q764m5 sus scrofa
27	562	48.0	749	1	STA1_MOUSE	P42225 mus musculu
28	540.5	46.1	754	2	O13131	O13131 oncorhynchu
29	533	45.5	528	2	Q8JFU8	Q8jfu8 brachydanio
30	532.5	45.4	108	2	Q704W6	Q704w6 bos taurus
31	532	45.4	749	2	O93598	O93598 brachydanio
32	532	45.4	749	2	Q6P943	Q6p943 brachydanio
33	531.5	45.3	718	2	Q801Y2	Q801y2 carassius a
34	520	44.4	748	1	STA4_HUMAN	Q14765 homo sapien
35	519.5	44.3	758	2	Q90Y17	Q90y17 tetraodon f
36	517.5	44.2	754	2	O13132	O13132 oncorhynchu
37	513	43.8	657	2	Q8AW24	Q8aw24 brachydanio
38	504.5	43.0	749	1	STA4_MOUSE	P42228 mus musculu
39	504	43.0	651	2	Q7ZZ53	Q7zz53 brachydanio
40	504	43.0	667	2	Q8AW20	Q8aw20 brachydanio
41	503.5	43.0	652	2	Q7ZZ77	Q7zz77 brachydanio
42	503	42.9	553	2	Q8JFU7	Q8jfu7 brachydanio
43	503	42.9	748	2	Q66HB2	Q66hb2 rattus norv
44	503	42.9	1153	2	Q8JFS5	Q8jfs5 brachydanio
45	491	41.9	737	2	Q90Y15	Q90y15 tetraodon f

ALIGNMENTS

RESULT 1
STA3_HUMAN
ID STA3_HUMAN STANDARD; PRT; 770 AA.
AC P40763; O14916; Q9BW54;
DT 01-FEB-1995 (Rel. 31, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Signal transducer and activator of transcription 3 (Acute-phase
DE response factor).
GN Name=STAT3; Synonyms=APRF;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Placenta;
RX MEDLINE=94208062; PubMed=7512451; DOI=10.1016/0092-8674(94)90235-6;
RA Akira S., Nishio Y., Inoue M., Wang X.-J., Wei S., Matsusaka T.,
RA Yoshida K., Sudo T., Naruto M., Kishimoto T.;
RT "Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-
RT related transcription factor involved in the gp130-mediated signaling
RT pathway.";
RL Cell 77:63-71(1994).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=98296260; PubMed=9630560; DOI=10.1016/S0378-1119(98)00185-1;
RA Della Pietra L., Bressan A., Pezzotti A., Serlupi-Crescenzi O.;
RT "Highly conserved amino-acid sequence between murine STAT3 and a
RT revised human STAT3 sequence.";
RL Gene 213:119-124(1998).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM DEL-701), AND VARIANT ILE-143.
RA Rieder M.J., Daniels R.L., da Ponte S.H., Hastings N.C., Ahearn M.O.,
RA Rajkumar N., Yi Q., Nickerson D.A.;
RT "SeattleSNPs. NHLBI HL66682 program for genomic applications, UW-
RT FHCRC, Seattle, WA (URL: http://pga.gs.washington.edu).";
RL Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND DEL-701).
RC TISSUE=Kidney, and Pancreas;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,

```
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   SEQUENCE OF 564-704 FROM N.A.
RC   TISSUE=Liver;
RA   Della Pietra L., Bressan A., Pezzotti A.R., Serlupi-Crescenzi O.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   PHOSPHORYLATION ON SERINE.
RX   MEDLINE=95215843; PubMed=7701321;
RA   Zhang X., Blenis J., Li H.-C., Schindler C., Chen-Kiang S.;
RT   "Requirement of serine phosphorylation for formation of STAT-promoter
RT   complexes.";
RL   Science 267:1990-1994(1995).
RN   [7]
RP   INTERACTION WITH NCOA1.
RX   PubMed=11773079; DOI=10.1074/jbc.M111486200;
RA   Giraud S., Bienvenu F., Avril S., Gascan H., Heery D.M., Coqueret O.;
RT   "Functional interaction of STAT3 transcription factor with the
RT   coactivator NcoA/SRC1a.";
RL   J. Biol. Chem. 277:8004-8011(2002).
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes.
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P40763-1; Sequence=Displayed;
CC       Name=Del-701;
CC         IsoId=P40763-2; Sequence=VSP_010474;
CC   -!- TISSUE SPECIFICITY: Heart, brain, placenta, lung, liver, skeletal
CC       muscle, kidney and pancreas.
CC   -!- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,
CC       LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation
CC       is important for the formation of stable DNA-binding STAT3
CC       homodimers and maximal transcriptional activity.
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L29277; AAA58374.1; -.
DR   EMBL; AJ012463; CAA10032.1; -.
DR   EMBL; AY572796; AAS66986.1; -.
DR   EMBL; BC000627; AAH00627.1; -.
DR   EMBL; BC014482; AAH14482.1; -.
DR   EMBL; AF029311; AAB84254.1; -.
DR   PIR; A54444; A54444.
DR   HSSP; P42227; 1BG1.
DR   TRANSFAC; T01493; -.
DR   Genew; HGNC:11364; STAT3.
DR   H-InvDB; HIX0013840; -.
DR   MIM; 102582; -.
DR   GO; GO:0005737; C:cytoplasm; TAS.
DR   GO; GO:0005634; C:nucleus; TAS.
DR   GO; GO:0005062; F:hematopoietin/interferon-class (D200-domain. . .; TAS.
DR   GO; GO:0003700; F:transcription factor activity; TAS.
DR   GO; GO:0006928; P:cell motility; TAS.
DR   GO; GO:0007259; P:JAK-STAT cascade; TAS.
DR   GO; GO:0000122; P:negative regulation of transcription from P. . .; TAS.
DR   GO; GO:0007399; P:neurogenesis; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
KW   Activator; Alternative splicing; DNA-binding; Nuclear protein;
KW   Phosphorylation; Polymorphism; SH2 domain; Transcription regulation.
FT   DOMAIN      580    670       SH2.
FT   MOD_RES     705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES     727    727       Phosphoserine (By similarity).
FT   VARSPLIC    701    701       Missing (in isoform Del-701).
FT                                /FTId=VSP_010474.
FT   VARIANT      32     32       Q -> K (in dbSNP:1803125).
FT                                /FTId=VAR_018683.
FT   VARIANT     143    143       M -> I.
FT                                /FTId=VAR_018679.
FT   CONFLICT    288    288       Q -> H (in Ref. 1).
FT   CONFLICT    460    460       P -> S (in Ref. 1).
FT   CONFLICT    548    548       K -> N (in Ref. 1).
FT   CONFLICT    561    561       F -> Y (in Ref. 1).
FT   CONFLICT    667    667       V -> L (in Ref. 1).
FT   CONFLICT    730    730       T -> A (in Ref. 1).
SQ   SEQUENCE   770 AA;  88067 MW;  6C00632211C8012D CRC64;

  Query Match            98.7%;  Score 1157;  DB 1;  Length 770;
  Best Local Similarity  98.7%;  Pred. No. 2.9e-65;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 2
STA3_MOUSE
ID   STA3_MOUSE     STANDARD;      PRT;   770 AA.
AC   P42227;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Signal transducer and activator of transcription 3 (Acute-phase
DE   response factor).
GN   Name=Stat3; Synonyms=Aprf;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A), AND SEQUENCE OF 154-158; 181-185
RP   AND 632-640.
RC   STRAIN=BALB/c; TISSUE=Liver;
RX   MEDLINE=94208062; PubMed=7512451; DOI=10.1016/0092-8674(94)90235-6;
```

```
RA   Akira S., Nishio Y., Inoue M., Wang X.J., Shi W., Matsusaka T.,
RA   Yoshida K., Sudo T., Naruto M., Kishimoto T.;
RT   "Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-
RT   related transcription factor involved in the gp130-mediated signaling
RT   pathway.";
RL   Cell 77:63-71(1994).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   TISSUE=Thymus;
RX   MEDLINE=94188718; PubMed=8140422;
RA   Zhong Z., Wen Z., Darnell J.E. Jr.;
RT   "Stat3: a STAT family member activated by tyrosine phosphorylation in
RT   response to epidermal growth factor and interleukin-6.";
RL   Science 264:95-98(1994).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   TISSUE=Brain;
RX   MEDLINE=95014185; PubMed=7523373;
RA   Raz R., Durbin J.E., Levy D.E.;
RT   "Acute phase response factor and additional members of the interferon-
RT   stimulated gene factor 3 family integrate diverse signals from
RT   cytokines, interferons, and growth factors.";
RL   J. Biol. Chem. 269:24391-24395(1994).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3B).
RC   STRAIN=BALB/c, and C57BL/6; TISSUE=Liver;
RX   MEDLINE=96016116; PubMed=7568080;
RA   Schaefer T.S., Sanders L.K., Nathans D.;
RT   "Cooperative transcriptional activity of Jun and Stat3 beta, a short
RT   form of Stat3.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:9097-9101(1995).
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   STRAIN=129/SvJ;
RX   PubMed=11161808; DOI=10.1006/geno.2000.6433;
RA   Miyoshi K., Cui Y., Riedlinger G., Lehoczky J., Zon L., Oka T.,
RA   Dewar K., Hennighausen L.;
RT   "Structure of the mouse stat 3/5 locus: evolution from Drosophila to
RT   zebrafish to mouse.";
RL   Genomics 71:150-155(2001).
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   STRAIN=C57BL/6J, and NOD/LtJ;
RA   Davoodi-Semiromi A., She J.-X.;
RT   "A mutant Stat5b with weaker DNA binding defines a key defective
RT   pathway in non-obese diabetic (NOD) mice.";
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A. (ISOFORM STAT3A).
RC   STRAIN=FVB/N; TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [8]
RP   PHOSPHORYLATION SITE SER-727, AND MUTAGENESIS.
RX   MEDLINE=95354205; PubMed=7543024; DOI=10.1016/0092-8674(95)90311-9;
RA   Wen Z., Zhong Z., Darnell J.E. Jr.;
RT   "Maximal activation of transcription by Stat1 and Stat3 requires both
RT   tyrosine and serine phosphorylation.";
RL   Cell 82:241-250(1995).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 136-716.
RX   MEDLINE=98334373; PubMed=9671298; DOI=10.1038/28101;
RA   Becker S., Groner B., Mueller C.W.;
RT   "Three-dimensional structure of the Stat3beta homodimer bound to
RT   DNA.";
RL   Nature 394:145-151(1998).
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes. STAT3B interacts with the N-terminal
CC       part of JUN to activate such promoters in a cooperative way.
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=Stat3A;
CC         IsoId=P42227-1; Sequence=Displayed;
CC       Name=Stat3B;
CC         IsoId=P42227-2; Sequence=VSP_006287;
CC       Name=Del-701;
CC         IsoId=P42227-3; Sequence=VSP_010475;
CC   -!- TISSUE SPECIFICITY: STAT3A is seen in the liver, spleen, and
CC       kidney. STAT3B is also detected in the liver, although in a much
CC       less abundant manner.
CC   -!- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,
CC       LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation
CC       is important for the formation of stable DNA-binding STAT3
CC       homodimers and maximal transcriptional activity (By similarity).
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L29278; AAA37254.1; -.
DR   EMBL; U06922; AAA19452.1; -.
DR   EMBL; U08378; AAA56668.1; -.
DR   EMBL; U30709; AAC52612.1; -.
DR   EMBL; AF246978; AAL59017.1; -.
DR   EMBL; AY299489; AAQ75418.1; -.
DR   EMBL; AY299490; AAQ75419.1; -.
DR   EMBL; BC003806; AAH03806.1; -.
DR   PIR; I49508; I49508.
DR   PDB; 1BG1; X-ray; A=1-722.
DR   TRANSFAC; T01574; -.
DR   MGD; MGI:103038; Stat3.
DR   GO; GO:0005737; C:cytoplasm; IDA.
DR   GO; GO:0005634; C:nucleus; IDA.
DR   GO; GO:0005886; C:plasma membrane; IDA.
DR   GO; GO:0003677; F:DNA binding; IDA.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   GO; GO:0016563; F:transcriptional activator activity; IDA.
DR   GO; GO:0007259; P:JAK-STAT cascade; IDA.
DR   GO; GO:0006357; P:regulation of transcription from PoI II pro. . .; IDA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
```

```
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
KW   3D-structure; Activator; Acute phase; Alternative splicing;
KW   Direct protein sequencing; DNA-binding; Nuclear protein;
KW   Phosphorylation; SH2 domain; Transcription regulation.
FT   DOMAIN      580    670       SH2.
FT   MOD_RES     705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES     727    727       Phosphoserine.
FT   VARSPLIC    716    770       TTCSNTIDLPMSPRTLDSLMQFGNNGEGAEPSAGGQFESLT
FT                                FDMDLTSECATSPM -> FIDAVWK (in isoform
FT                                Stat3B).
FT                                /FTId=VSP_006287.
FT   VARSPLIC    701    701       Missing (in isoform Del-701).
FT                                /FTId=VSP_010475.
FT   MUTAGEN     727    727       S->A: Decreased transcriptional
FT                                activation.
FT   CONFLICT     16     16       E -> K (in Ref. 2).
FT   CONFLICT     25     25       S -> T (in Ref. 2 and 4).
FT   CONFLICT    394    394       M -> I (in Ref. 1).
FT   HELIX       139    180
FT   TURN        181    182
FT   TURN        197    198
FT   HELIX       199    237
FT   TURN        238    238
FT   HELIX       239    251
FT   TURN        252    253
FT   HELIX       261    290
FT   TURN        294    295
FT   TURN        297    301
FT   HELIX       302    320
FT   STRAND      321    328
FT   TURN        330    331
FT   TURN        333    334
FT   TURN        336    337
FT   STRAND      338    340
FT   TURN        341    342
FT   STRAND      345    351

  Query Match             98.7%;  Score 1157;  DB 1;  Length 770;
  Best Local Similarity   98.7%;  Pred. No. 2.9e-65;
  Matches  226;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 3
STA3_RAT
ID   STA3_RAT       STANDARD;      PRT;   770 AA.
AC   P52631;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Signal transducer and activator of transcription 3.
GN   Name=Stat3;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=96102059; PubMed=8530402; DOI=10.1074/jbc.270.50.29998;
RA   Ripperger J.A., Fritz S., Richter K., Hocke G.M., Lottspeich F.,
RA   Fey G.H.;
RT   "Transcription factors Stat3 and Stat5b are present in rat liver
RT   nuclei late in an acute phase response and bind interleukin-6 response
RT   elements.";
RL   J. Biol. Chem. 270:29998-30006(1995).
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes.
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation (By similarity).
CC   -!- PTM: Tyrosine phosphorylated in response to IL-6, IL-11, CNTF,
CC       LIF, CSF-1, EGF, PDGF, IFN-alpha and OSM. Serine phosphorylation
CC       is important for the formation of stable DNA-binding STAT3
CC       homodimers and maximal transcriptional activity (By similarity).
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X91810; CAA62920.1; -.
DR   HSSP; P42227; 1BG1.
DR   RGD; 3772; Stat3.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
KW   Activator; DNA-binding; Nuclear protein; Phosphorylation; SH2 domain;
KW   Transcription regulation.
FT   DOMAIN      580    670       SH2.
FT   MOD_RES     705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES     727    727       Phosphoserine (By similarity).
SQ   SEQUENCE   770 AA;  88039 MW;  D74A0C76954754ED CRC64;

  Query Match             98.5%;  Score 1154;  DB 1;  Length 770;
  Best Local Similarity   98.3%;  Pred. No. 4.5e-65;
  Matches  225;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |:||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 4
```

```
STA3_BOVIN
ID   STA3_BOVIN     STANDARD;      PRT;   770 AA.
AC   P61635;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Signal transducer and activator of transcription 3.
GN   Name=STAT3;
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Mammary gland;
RA   Seyfert H.M., Wheeler T.T., Moolenaar A., Pitra C.;
RT   "The STAT5B-encoding gene was flipped across the STAT3/STAT5A-locus
RT   during ruminant evolution.";
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Transcription factor that binds to the interleukin-6
CC       (IL-6)-responsive elements identified in the promoters of various
CC       acute-phase protein genes (By similarity).
CC   -!- PATHWAY: Involved in the gp130-mediated signaling pathway.
CC   -!- SUBUNIT: Forms a homodimer or a heterodimer with a related family
CC       member (at least STAT1). Interacts with NCOA1 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic; translocated into the nucleus
CC       in response to phosphorylation (By similarity).
CC   -!- SIMILARITY: Belongs to the transcription factor STAT family.
CC   -!- SIMILARITY: Contains 1 SH2 domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ620655; CAF06182.1; -.
DR   PROSITE; PS50001; SH2; 1.
KW   Activator; DNA-binding; Nuclear protein; Phosphorylation; SH2 domain;
KW   Transcription regulation.
FT   DOMAIN      580    670       SH2.
FT   MOD_RES     705    705       Phosphotyrosine (by JAK) (By similarity).
FT   MOD_RES     727    727       Phosphoserine (By similarity).
SQ   SEQUENCE   770 AA;  87974 MW;  9CEB147C73E83274 CRC64;

  Query Match             98.1%;  Score 1150;  DB 1;  Length 770;
  Best Local Similarity   98.3%;  Pred. No. 8.1e-65;
  Matches  225;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              |||||||||||||||||||| |||||||||||||| ||||||||||||||||||||||||
Db        130 NHPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLLDDFDFNYKTLKSQGDMQDLN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 358


RESULT 5
Q6DV79
ID   Q6DV79      PRELIMINARY;      PRT;   771 AA.
AC   Q6DV79;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Signal transducer and activator of transcription 3.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Zhou G.Y., Leung F.C.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY641397; AAT64887.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   771 AA;  88174 MW;  71AC855C5DEC03E2 CRC64;

  Query Match             96.6%;  Score 1132;  DB 2;  Length 771;
  Best Local Similarity   96.9%;  Pred. No. 1.1e-63;
  Matches  221;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 HPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNG 61
              |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        131 HPTAAVVTEKQQMLEQHLQDVRKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLNG 190

Qy         62 NNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEIA 121
              |||||||||||||||||||||||| |||||||||||||| | |||||||||| :||
Db        191 NNQSVTRQKMQQLEQMLTALDQMRRGIVSELAGLLSAMEYVQKMLADEELADWKRRQQIA 250

Qy        122 CIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERIV 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 CIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERIV 310

Qy        182 ELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||| ||||||||||||||||
Db        311 ELFRNLMKSAFVVERQPCMPMHPDRPLVFKTGVQFTTKVRLLVKFPEL 358


RESULT 6
Q9PVX8
ID   Q9PVX8      PRELIMINARY;      PRT;   769 AA.
AC   Q9PVX8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat 3.
GN   Name=stat 3;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20107399; PubMed=10642787; DOI=10.1006/dbio.1999.9518;
RA   Nishinakamura R., Matsumoto Y., Matsuda T., Ariizumi T., Heike T.,
RA   Asashima M., Yokota T.;
RT   "Activation of Stat3 by cytokine receptor gp130 ventralizes Xenopus
RT   embryos independent of BMP-4.";
RL   Dev. Biol. 216:481-490(1999).
DR   EMBL; AB017701; BAA86061.1; -.
```

```
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   769 AA;  87974 MW;  0905C03263303069 CRC64;

  Query Match          93.8%;  Score 1099;  DB 2;  Length 769;
  Best Local Similarity   92.6%;  Pred. No. 1.4e-61;
  Matches  212;  Conservative    9;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              :|| |||||||||||||||| ||:|||||||||||||||||||||||||| |: :||
Db        130 SHPNAAVVTEKQQMLEQHLQDVRKKVQDLEQKMKVVENLQDDFDFNYKTLKSQSDLSELN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||:||:|:|:|| |||||||||||||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQLRRTIISDLASLLSAMEYVQKTLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQIRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||| ||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTNKVRLLVKFPEL 358


RESULT 7
Q7ZXK3
ID   Q7ZXK3      PRELIMINARY;      PRT;   766 AA.
AC   Q7ZXK3;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat3-A protein.
GN   Name=stat3-A;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA   Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA   Richardson P.;
RT   "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT   initiative.";
RL   Dev. Dyn. 225:384-391(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC044717; AAH44717.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   766 AA;  87599 MW;  31018A3321CCEB9C CRC64;

  Query Match          93.5%;  Score 1096;  DB 2;  Length 766;
  Best Local Similarity   91.7%;  Pred. No. 2.1e-61;
  Matches  210;  Conservative   11;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDMQDLN 60
              :|| |||:|||||||||||| ||:|||||||||||||||||||||||| |: :||
Db        130 SHPNAAVITEKQQMLEQHLQDVRKKVQDLEQKMKVVENLQDDFDFNYKTLKSQSDLSELN 189

Qy         61 GNNQSVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPEI 120
              |||||||||||||||||||:||:|:|:|| |||||||||| |||||||||| :|
Db        190 GNNQSVTRQKMQQLEQMLTALDQLRRTIISDLASLLSAMEYVQKNLTDEELADWKRRQQI 249

Qy        121 ACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEERI 180
              |||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db        250 ACIGGPPNICLDRLENWITSLAESQLQTRQQIRKLEELQQKVSYKGDPIVQHRPMLEERI 309

Qy        181 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              |||||||||||||||||||||||||||||||||||| ||||||||||
Db        310 VELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTNKVRLLVKFPEL 358


RESULT 8
Q7ZTS5
ID   Q7ZTS5      PRELIMINARY;      PRT;   414 AA.
AC   Q7ZTS5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat3 protein.
GN   Name=stat3;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB; TISSUE=Whole body;
```

```
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AB; TISSUE=Whole body;
RA   Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC045276; AAH45276.1; -.
DR   HSSP; P42227; 1BG1.
DR   ZFIN; ZDB-GENE-980526-68; stat3.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
SQ   SEQUENCE   414 AA;  48253 MW;  0FFD1B509B7526BD CRC64;

  Query Match             86.3%;  Score 1011;  DB 2;  Length 414;
  Best Local Similarity   84.8%;  Pred. No. 2.7e-56;
  Matches  195;  Conservative   20;  Mismatches   13;  Indels    2;  Gaps    2;

Qy          2 HPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLN 60
              |||  |||||||:|| :||| ||||||:|||||::|||||||||||||||||:: ||||
Db        130 HPTGTVVTEKQQILEHNLQDIRKRVQDMEQKMKMLENLQDDFDFNYKTLKSAGELSQDLN 189

Qy         61 GNNQ-SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPE 119
              ||:| : ||||| ||||||:||||:|| ||:|:|||||||::||| |||||||||||| :
Db        190 GNSQAAATRQKMSQLEQMLSALDQLRRQIVTEMAGLLSAMDFVQKNLTDEELADWKRRQQ 249

Qy        120 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 179
              |||||||||||||| |||||||||| ||||:||||||||||||:|||| |||:
Db        250 IACIGGPPNICLDRLETWITSLAESQLQIRQQIRKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy        180 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              ||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        310 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 359


RESULT 9
Q6DVF3
ID   Q6DVF3      PRELIMINARY;      PRT;   765 AA.
AC   Q6DVF3;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Signal transducer and activation of transcription factor 3.
OS   Oryzias latipes (Medaka fish) (Japanese ricefish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Liu R., Hong Y.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY639947; AAT64912.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   765 AA;  87566 MW;  F5D01408748EC703 CRC64;

  Query Match             86.3%;  Score 1011;  DB 2;  Length 765;
  Best Local Similarity   84.8%;  Pred. No. 5.2e-56;
  Matches  195;  Conservative   19;  Mismatches   14;  Indels    2;  Gaps    2;

Qy          2 HPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLN 60
              |||  |||||||:|| :||| ||||||:|||||::|||||||||||||||||:: ||||
Db        130 HPTGTVVTEKQQILEHNLQDIRKRVQDMEQKMKMLENLQDDFDFNYKTLKSQGELNQDLN 189

Qy         61 GNNQ-SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPE 119
              ||:| : ||||| ||||||:||||:|| ||:|:|||:||||:|||| |||||||||||| :
Db        190 GNSQAAATRQKMAQLEQMLSALDQLRRQIVTEMGGLLTAMDYVQKNLTDEELADWKRRQQ 249

Qy        120 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 179
              |||||||||||||| |||||||||| ||||:||||||||||||:|||| |||:
Db        250 IACIGGPPNICLDRLETWITSLAESQLQIRQQIKKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy        180 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              ||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        310 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTNKVRLLVKFPEL 359


RESULT 10
Q6GUE7
ID   Q6GUE7      PRELIMINARY;      PRT;   785 AA.
AC   Q6GUE7;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Signal transducer and activator of transcription 3 isoform 1.
OS   Oryzias latipes (Medaka fish) (Japanese ricefish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Liu R., Hong Y.;
RL   Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY641434; AAT46364.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
```

```
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   785 AA;  89643 MW;  81F231BDE27DE938 CRC64;

  Query Match             86.3%;  Score 1011;  DB 2;  Length 785;
  Best Local Similarity   84.8%;  Pred. No. 5.3e-56;
  Matches  195;  Conservative   19;  Mismatches   14;  Indels    2;  Gaps    2;

Qy          2 HPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLN 60
              |||  |||||||:|| :||| ||||||:|||||::|||||||||||||||:: ||||
Db        130 HPTGTVVTEKQQILEHNLQDIRKRVQDMEQKMKMLENLQDDFDFNYKTLKSQGELNQDLN 189

Qy         61 GNNQ-SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPE 119
              ||:| : ||||| |||||:||||:|| ||:|: |||:||:|||| |||||||||||| :
Db        190 GNSQAAATRQKMAQLEQMLSALDQLRRQIVTEMGGLLTAMDYVQKNLTDEELADWKRRQQ 249

Qy        120 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 179
              |||||||||||||||| |||||||||| ||||||||||||||||||:|||| |||:
Db        250 IACIGGPPNICLDRLETWITSLAESQLQIRQQIKKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy        180 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              ||:|||||||||||||||||||||||||||||||| ||||||||||
Db        310 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTNKVRLLVKFPEL 359


RESULT 11
Q6NV46
ID   Q6NV46      PRELIMINARY;      PRT;   786 AA.
AC   Q6NV46;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Stat3 protein.
GN   Name=stat3;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RA   Strausberg R.;
RL   Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC068320; AAH68320.1; -.
DR   ZFIN; ZDB-GENE-980526-68; stat3.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   786 AA;  90039 MW;  FC7371D0B0E5447E CRC64;

  Query Match             86.3%;  Score 1011;  DB 2;  Length 786;
  Best Local Similarity   84.8%;  Pred. No. 5.4e-56;
  Matches  195;  Conservative   20;  Mismatches   13;  Indels    2;  Gaps    2;

Qy          2 HPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLN 60
              |||  |||||||:|| :||| ||||||:|||||::|||||||||||||||:: ||||
Db        130 HPTGTVVTEKQQILEHNLQDIRKRVQDMEQKMKMLENLQDDFDFNYKTLKSAGELSQDLN 189

Qy         61 GNNQ-SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPE 119
              ||:| : ||||| |||||:||||:|| ||:|: |||:||:|||| |||||||||||| :
Db        190 GNSQAAATRQKMSQLEQMLSALDQLRRQIVTEMAGLLLSAMDFVQKNLTDEELADWKRRQQ 249

Qy        120 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 179
              |||||||||||||||| |||||||||| ||||:||||||||||||:|||| |||:
Db        250 IACIGGPPNICLDRLETWITSLAESQLQIRQQIRKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy        180 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              ||:|||||||||||||||||||||||||||||||| ||||||||||
Db        310 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 359


RESULT 12
O93599
ID   O93599      PRELIMINARY;      PRT;   806 AA.
AC   O93599;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Transcription factor.
GN   Name=stat3;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Oates A.C.;
RL   Thesis (1998), University of Melbourne, Australia.
DR   EMBL; AJ005693; CAA06677.1; -.
DR   HSSP; P42227; 1BG1.
DR   ZFIN; ZDB-GENE-980526-68; stat3.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   806 AA;  92151 MW;  74BC4EA401C3C942 CRC64;

  Query Match             86.3%;  Score 1011;  DB 2;  Length 806;
  Best Local Similarity   84.8%;  Pred. No. 5.5e-56;
```

```
  Matches  195;  Conservative   20;  Mismatches   13;  Indels    2;  Gaps    2;

Qy          2 HPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLN 60
              |||  |||||||:|| :||| ||||||:|||||::|||||||||||||||| |:: ||||
Db        130 HPTGTVVTEKQQILEHNLQDIRKRVQDMEQKMKMLENLQDDFDFNYKTLKSAGELSQDLN 189

Qy         61 GNNQ-SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPE 119
              ||:| : ||||| |||||:||||:|| ||:|:|||||||::||| ||||||||||| :
Db        190 GNSQAAATRQKMSQLEQMLSALDQLRRQIVTEMAGLLSAMDFVQKNLTDEELADWKRRQQ 249

Qy        120 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 179
              |||||||||||||||| |||||||||||| ||||:||||||||||||||:|||| |||:
Db        250 IACIGGPPNICLDRLETWITSLAESQLQIRQQIRKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy        180 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              ||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 359


RESULT 13
Q90Y16
ID   Q90Y16      PRELIMINARY;      PRT;   764 AA.
AC   Q90Y16;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   STAT3.
OS   Tetraodon fluviatilis (Puffer fish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=47145;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sung S.C., Fan T.J., Leu J.H., Chou C.M., Huang C.J.;
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF307106; AAL09415.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
DR   SMART; SM00252; SH2; 1.
DR   PROSITE; PS50001; SH2; 1.
SQ   SEQUENCE   764 AA;  87411 MW;  E661FFE18BEFD8BE CRC64;

  Query Match             84.2%;  Score 987;  DB 2;  Length 764;
  Best Local Similarity   83.5%;  Pred. No. 1.7e-54;
  Matches  192;  Conservative   20;  Mismatches   16;  Indels    2;  Gaps    2;

Qy          2 HPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDLN 60
              :|:  |||||||||| :||| ||||||:|||||::|||||||||||||||:: ||||
Db        130 NPSGTVVTEKQQMLELNLQDIRKRVQDMEQKMKMLENLQDDFDFNYKTLKSQGELSQDLN 189

Qy         61 GNNQ-SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRPE 119
              ||:| |  |||| ||||||:||||:|| ||:|: |||:||:|||| ||:|||||||| :
Db        190 GNSQASAARQKMAQLEQMLSALDQLRRQIVTEMGGLLTAMDYVQKNLTNEELADWKRRQQ 249

Qy        120 IACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEER 179
              |||||||||||||||| |||||||||||| |||||||||||||||||||:|||| |||:
Db        250 IACIGGPPNICLDRLETWITSLAESQLQIRQQIKKLEELQQKVSYKGDPIIQHRPALEEK 309

Qy        180 IVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              ||:|||||||||||||||||||||||||||||||| |||||||| ||
Db        310 IVDLFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTNKVRLLVKFQEL 359


RESULT 14
O13133
ID   O13133      PRELIMINARY;      PRT;   767 AA.
AC   O13133;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stat3.
GN   Name=rbtStat3;
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8022;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Johnson M.C., Mourich D.V., Leong J.C.;
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U60333; AAB60926.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR000980; SH2.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF00017; SH2; 1.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
DR   Pfam; PF02865; STAT_int; 1.
SQ   SEQUENCE   767 AA;  87816 MW;  FED97740C74C3798 CRC64;

  Query Match             83.4%;  Score 977;  DB 2;  Length 767;
  Best Local Similarity   81.4%;  Pred. No. 7.4e-54;
  Matches  188;  Conservative   24;  Mismatches   17;  Indels    2;  Gaps    2;

Qy          1 NHPTAAVVTEKQQMLEQHLQDARKRVQDLEQKMKVVENLQDDFDFNYKTLKSQGDM-QDL 59
              :||:  |||||||:|| :||| ||||||:|||||::|||||||||:|||||||:: ||:
Db        129 SHPSGTVVTEKQQILEHNLQDIRKRVQDMEQKMKMLENLQDDFDFHYKTLKSQGELSQDM 188

Qy         60 NGNNQ-SVTRQKMQQLEQMLTALDQMRRSIVSELAGLLSAMEYVQKTLTDEELADWKRRP 118
              |||:| : ||||| ||||||:||||:|| ||:|:|||||||::||| |||:|||||||
Db        189 NGNSQAAATRQKMSQLEQMLSALDQLRRQIVTEMAGLLSAMDFVQKNLTDDELADWKRRQ 248

Qy        119 EIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQKVSYKGDPIVQHRPMLEE 178
              :|||||||| ||||||| ||||| | ||| |||||||||||||||||||:|||| |||
Db        249 QIACIGGPPKICLDRLETWITSLGEIQLQIRQQIKKLEELQQKVSYKGDPIIQHRPALEE 308

Qy        179 RIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVRLLVKFPEL 229
              :||:|||||||||||| |||||||||||||||||| ||||||||||
Db        309 KIVDLFRNLMKSAFVVERQRCMPMHPDRPLVIKTGVQFTNKVRLLVKFPEL 359


RESULT 15
Q9N145
ID   Q9N145      PRELIMINARY;      PRT;   163 AA.
AC   Q9N145;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Signal transducer and activator of transcription (Fragment).
GN   Name=STAT3;
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Arredondo J.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; AF227560; AAF73401.1; -.
DR   HSSP; P42227; 1BG1.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0004871; F:signal transducer activity; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   GO; GO:0007165; P:signal transduction; IEA.
DR   InterPro; IPR008967; P53_like_DNA_bnd.
DR   InterPro; IPR001217; STAT.
DR   Pfam; PF01017; STAT_alpha; 1.
DR   Pfam; PF02864; STAT_bind; 1.
FT   NON_TER       1      1
FT   NON_TER     163    163
SQ   SEQUENCE   163 AA;  18855 MW;  DD3F11E7F5B9878C CRC64;

  Query Match             56.0%;  Score 656;  DB 2;  Length 163;
  Best Local Similarity   98.4%;  Pred. No. 3.1e-34;
  Matches  126;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        102 VQKTLTDEELADWKRRPEIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQK 161
              ||||||||||||||| :||||||||||||||||||||||||||||||||||||||||||
Db          1 VQKTLTDEELADWKRRQQIACIGGPPNICLDRLENWITSLAESQLQTRQQIKKLEELQQK 60

Qy        162 VSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVR 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VSYKGDPIVQHRPMLEERIVELFRNLMKSAFVVERQPCMPMHPDRPLVIKTGVQFTTKVR 120

Qy        222 LLVKFPEL 229
              ||||||||
Db        121 LLVKFPEL 128


Search completed: May 25, 2005, 17:43:45
Job time : 96.284 secs
```